STIC Search Report
Biotech-Chem Library

STIC Database Tracking Number: 115605

To: James Schultz **Location:** rem/2d18/2c18 **Art Unit:** 1635 Monday, March 08, 2004 **Case Serial Number:** 09/646651	**From:** Beverly Shears **Location:** Remsen Bldg. RM 1A54 **Phone:** 571-272-2528 beverly.shears@uspto.gov

Search Notes

U.S. DEPARTMENT OF COMMERCE
Patent and Trademark Office

SEARCH REQUEST FORM

Requestor's Name: _____ Serial Number: _____

Date: _____ Phone: _____ Art Unit: _____

Search Topic:
Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevent citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevent claim(s).

STAFF USE ONLY

Date completed: 03-05-04
Searcher: Beverly C 2528
Terminal time: 20
Elapsed time: _____
CPU time: _____
Total time: 25
Number of Searches: _____
Number of Databases: 1

Search Site
_____ STIC
_____ CM-1
_____ Pre-S

Type of Search
_____ N.A. Sequence
_____ A.A. Sequence
_____ Structure
_____ Bibliographic

Vendors
_____ IG
_____ STN
_____ Dialog
_____ APS
_____ Geninfo
_____ SDC
_____ DARC/Questel
✓ Other CGN

09/646651
Seq. IDs 1 & 2 w/ Intros

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

M protein - protein search, using sw model

un on: March 5, 2004, 10:31:00 ; Search time 182 Seconds
(without alignments)
488.027 Million cell updates/sec

itle: US-09-646-651C-1
erfect score: 467
equence: 1 TKLEDHLEGIINIFHQYSVR.....EFVVLVTDVLITAHDNIHKE 91

coring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

earched: 6019581 seqs, 976053577 residues

otal number of hits satisfying chosen parameters: 6019581

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US080_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US097A_COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US097B_COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
24: /cgn2_6/ptodata/2/paa/US099A_COMB.pep:*
25: /cgn2_6/ptodata/2/paa/US099B_COMB.pep:*
26: /cgn2_6/ptodata/2/paa/US100_COMB.pep:*
27: /cgn2_6/ptodata/2/paa/US101_COMB.pep:*
28: /cgn2_6/ptodata/2/paa/US102_COMB.pep:*
29: /cgn2_6/ptodata/2/paa/US103_COMB.pep:*
30: /cgn2_6/ptodata/2/paa/US104_COMB.pep:*
31: /cgn2_6/ptodata/2/paa/US106_COMB.pep:*
32: /cgn2_6/ptodata/2/paa/US107_COMB.pep:*
33: /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	467	100.0	91	1	US-08-794-000-2	Sequence 2, Appli
2	467	100.0	91	11	US-08-759-913-5	Sequence 5, Appli
3	467	100.0	91	20	US-09-646-651C-1	Sequence 1, Appli
4	467	100.0	91	22	US-09-791-537-101828	Sequence 101828,
5	455	97.4	91	11	US-08-761-289-9	Sequence 9, Appli
6	455	97.4	91	20	US-09-646-264A-1	Sequence 1, Appli
7	455	97.4	91	20	US-09-646-651A-1	Sequence 1, Appli
8	455	97.4	91	20	US-09-646-651B-1	Sequence 1, Appli
9	371	79.4	92	22	US-09-791-537-99618	Sequence 99618, A
10	366	78.4	90	15	US-09-167-705-3	Sequence 3, Appli
11	366	78.4	90	23	US-09-872-185-11	Sequence 11, Appl
12	366	78.4	90	23	US-09-872-185-12	Sequence 12, Appl
13	366	78.4	90	23	US-09-872-185B-11	Sequence 11, Appl
14	366	78.4	90	23	US-09-872-185B-12	Sequence 12, Appl
15	366	78.4	90	31	US-10-666-513-3	Sequence 3, Appli
16	332	71.1	91	22	US-09-791-537-132106	Sequence 132106,
17	332	71.1	91	26	US-10-030-937-21	Sequence 21, Appl
18	332	71.1	92	1	PCT-US95-16871-2	Sequence 2, Appli
19	332	71.1	92	11	US-08-759-913-1	Sequence 1, Appli
20	332	71.1	92	11	US-08-761-289-2	Sequence 2, Appli
21	332	71.1	92	16	US-09-227-854-2	Sequence 2, Appli
22	332	71.1	92	22	US-09-791-537-137536	Sequence 137536,
23	332	71.1	92	25	US-09-958-053-24	Sequence 24, Appl
24	332	71.1	92	26	US-10-030-937-19	Sequence 19, Appl
25	332	71.1	92	26	US-10-030-937-20	Sequence 20, Appl
26	332	71.1	92	26	US-10-030-937-23	Sequence 23, Appl
27	332	71.1	92	26	US-10-077-600-2	Sequence 2, Appli
28	332	71.1	92	32	US-10-755-889-334	Sequence 334, App
29	332	71.1	92	33	US-60-440-068-334	Sequence 334, App
30	332	71.1	92	33	US-60-452-680-12625	Sequence 12625, A
31	332	71.1	92	33	US-60-453-050-7882	Sequence 7882, Ap
32	332	71.1	92	33	US-60-453-135-7882	Sequence 7882, Ap
33	332	71.1	92	33	US-60-455-444-4409	Sequence 4409, Ap
34	332	71.1	92	33	US-60-465-241-4409	Sequence 4409, Ap
35	332	71.1	92	33	US-60-466-412-7882	Sequence 7882, Ap
36	332	71.1	92	33	US-60-469-757-334	Sequence 334, App
37	332	71.1	95	22	US-09-760-443-1495	Sequence 1495, Ap
38	332	71.1	95	22	US-09-760-457-432	Sequence 432, App
39	332	71.1	95	22	US-09-760-484-588	Sequence 588, App
40	332	71.1	95	28	US-10-212-054-1495	Sequence 1495, Ap
41	332	71.1	95	28	US-10-217-527-432	Sequence 432, App
42	306	65.5	95	1	PCT-US01-08631-57941	Sequence 57941, A
43	302	64.7	81	22	US-09-791-537-13830	Sequence 13830, A
44	289	61.9	70	22	US-09-791-537-138681	Sequence 138681,
45	271	58.0	139	22	US-09-760-484-450	Sequence 450, App

ALIGNMENTS

```
RESULT 1
US-08-794-000-2
; Sequence 2, Application PC/TDE9601337
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Metal-Containing Ribonucleotide Polypeptides
;    NUMBER OF SEQUENCES:  4
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;    CURRENT APPLICATION DATA:  APPLICATION NUMBER: US 08/794,000
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/DE96/01337
;      FILING DATE:  17-JUL-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 195 25 992.0
;      FILING DATE:  17-JUL-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 195 30 500.0
;      FILING DATE:  18-AUG-1995
```

```
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  91 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-794-000-2

  Query Match           100.0%;  Score 467;  DB 1;  Length 91;
  Best Local Similarity  100.0%;  Pred. No. 1.8e-44;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              |||||||||||||||||||||||||||||||
Db         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91


RESULT 2
US-08-759-913-5
; Sequence 5, Application US/08759913
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Goli, Surya K.
,    APPLICANT:  Hillman, Jennifer L.
,    TITLE OF INVENTION:  NOVEL HUMAN S100 PROTEINS
,    NUMBER OF SEQUENCES:  8
,    CORRESPONDENCE ADDRESS:
,      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
,      STREET:  3174 Porter Drive
,      CITY:  Palo Alto
,      STATE:  CA
,      COUNTRY:  US
,      ZIP:  94304
,    COMPUTER READABLE FORM:
,      MEDIUM TYPE:  Diskette
,      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/759,913
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
,      APPLICATION NUMBER:
,      FILING DATE:
,    ATTORNEY/AGENT INFORMATION:
,      NAME:  Billings, Lucy J.
,      REGISTRATION NUMBER:  36,749
,      REFERENCE/DOCKET NUMBER:  PF-0172 US
,    TELECOMMUNICATION INFORMATION:
,      TELEPHONE:  415-855-0555
,      TELEFAX:  415-845-4166
,  INFORMATION FOR SEQ ID NO:  5:
     SEQUENCE CHARACTERISTICS:
,      LENGTH:  91 amino acids
,      TYPE:  amino acid
,      STRANDEDNESS:  single
,      TOPOLOGY:  linear
,    MOLECULE TYPE:  peptide
,    IMMEDIATE SOURCE:
,      LIBRARY:  GenBank
,      CLONE:  461678
US-08-759-913-5

  Query Match           100.0%;  Score 467;  DB 11;  Length 91;
  Best Local Similarity  100.0%;  Pred. No. 1.8e-44;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              |||||||||||||||||||||||||||||||
Db         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91


RESULT 3
US-09-646-651C-1
; Sequence 1, Application US/09646651C
; GENERAL INFORMATION:
;  APPLICANT: Kiesewetter, Stefan
;  APPLICANT:  Kuhn, Eckehard
;  APPLICANT:  Koch-Pelster, Brigitte
;  APPLICANT:  Brunner, Herwig
;  TITLE OF INVENTION: METAL-CONTAINING RIBONUCLEOTIDE POLYPEPTIDES
;  FILE REFERENCE: 206579
;  CURRENT APPLICATION NUMBER: US/09/646,651C
;  CURRENT FILING DATE:  2001-01-16
;  PRIOR APPLICATION NUMBER: PCT/EP98/07722
;  PRIOR FILING DATE: 1998-11-30
;  PRIOR APPLICATION NUMBER: DE 198 11 047.2
;  PRIOR FILING DATE: 1998-03-13
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;    LENGTH: 91
;    TYPE: PRT
;    ORGANISM: Unknown
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: ()..()
;    OTHER INFORMATION: Angiotropin-related protein
US-09-646-651C-1

  Query Match           100.0%;  Score 467;  DB 20;  Length 91;
  Best Local Similarity  100.0%;  Pred. No. 1.8e-44;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              |||||||||||||||||||||||||||||||
Db         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91


RESULT 4
US-09-791-537-101828
; Sequence 101828, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 101828
;    LENGTH: 91
;    TYPE: PRT
;    ORGANISM: Sus scrofa
US-09-791-537-101828

  Query Match           100.0%;  Score 467;  DB 22;  Length 91;
  Best Local Similarity  100.0%;  Pred. No. 1.8e-44;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
y          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60

y         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
             |||||||||||||||||||||||||||||||
b         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91


ESULT 5
S-08-761-289-9
 Sequence 9, Application US/08761289
  GENERAL INFORMATION:
    APPLICANT:  Ni, J., ET AL.
    TITLE OF INVENTION:  Human Chemotactic Cytokine I
    NUMBER OF SEQUENCES:  10
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN,
      ADDRESSEE:  CECCHI, STEWART & OLSTEIN
      STREET:  6 BECKER FARM ROAD
      CITY:  ROSELAND
      STATE:  NEW JERSEY
      COUNTRY:  USA
      ZIP:  07068
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  3.5 INCH DISKETTE
      COMPUTER:  IBM PS/2
      OPERATING SYSTEM:  MS-DOS
      SOFTWARE:  WORD PERFECT 5.1
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/761,289
      FILING DATE:  December 6, 1996
      CLASSIFICATION:  435
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  60/008387
      FILING DATE:  December 8, 1995
    ATTORNEY/AGENT INFORMATION:
      NAME:  MULLINS, J.G.
      REGISTRATION NUMBER:  33,073
      REFERENCE/DOCKET NUMBER:  325800-506 (PF210)
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  201-994-1700
      TELEFAX:  201-994-1744
  INFORMATION FOR SEQ ID NO:  9:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  91 AMINO ACIDS
      TYPE:  AMINO ACID
      STRANDEDNESS:
      TOPOLOGY:  LINEAR
    MOLECULE TYPE:  PEPTIDE
S-08-761-289-9

 Query Match             97.4%;  Score 455;  DB 11;  Length 91;
 Best Local Similarity   97.8%;  Pred. No. 4.1e-43;
 Matches   89;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

y          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
b          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLIYKELPNTLKNTKDQGTIDKIFQNL 60

y         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
             :|||||||||||||||||||||||||||||||
b         61 NANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91


ESULT 6
S-09-646-264A-1
 Sequence 1, Application US/09646264A
 GENERAL INFORMATION:
  APPLICANT: Kiesewetter, Stefan
  APPLICANT:  Seibt, Jorg-Volker
  APPLICANT:  Noll, Michaela
  TITLE OF INVENTION: METAL-CONTAINING RIBONUCLEOTIDE POLYPEPTIDES
  FILE REFERENCE: 206604
  CURRENT APPLICATION NUMBER: US/09/646,264A
  CURRENT FILING DATE:  2001-05-07
  PRIOR APPLICATION NUMBER: PCT/EP99/00834
  PRIOR FILING DATE: 1999-02-09
  PRIOR APPLICATION NUMBER: DE19810998.9
  PRIOR FILING DATE: 1998-03-13
  NUMBER OF SEQ ID NOS: 6
  SOFTWARE: PatentIn version 3.0
 SEQ ID NO 1
    LENGTH: 91
    TYPE: PRT
    ORGANISM: Unknown
    FEATURE:
    NAME/KEY: misc_feature
    LOCATION: ()..()
    OTHER INFORMATION: Angiotropin-related protein
US-09-646-264A-1

  Query Match             97.4%;  Score 455;  DB 20;  Length 91;
  Best Local Similarity   97.8%;  Pred. No. 4.1e-43;
  Matches   89;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
             |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db         1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKYLITKELPNTLKNTKDQGTIDKIFYNL 60

Qy        61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
             |||||||||||||||||||||||||||||||
Db        61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91


RESULT 7
US-09-646-651A-1
; Sequence 1, Application US/09646651A
; GENERAL INFORMATION:
;  APPLICANT: Kiesewetter, Stefan
;  APPLICANT:  Kuhn, Eckehard
;  APPLICANT:  Koch-Pelster, Brigitte
;  APPLICANT:  Brunner, Herwig
;  TITLE OF INVENTION: METAL-CONTAINING RIBONUCLEOTIDE POLYPEPTIDES
;  FILE REFERENCE: 206579
;  CURRENT APPLICATION NUMBER: US/09/646,651A
;  CURRENT FILING DATE:  2000-09-13
;  PRIOR APPLICATION NUMBER: PCT/EP98/07722
;  PRIOR FILING DATE: 1998-11-30
;  PRIOR APPLICATION NUMBER: DE 198 11 047.2
;  PRIOR FILING DATE: 1998-03-13
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;    LENGTH: 91
;    TYPE: PRT
;    ORGANISM: Unknown
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: ()..()
;    OTHER INFORMATION: Angiotropin-related protein
US-09-646-651A-1

  Query Match             97.4%;  Score 455;  DB 20;  Length 91;
  Best Local Similarity   97.8%;  Pred. No. 4.1e-43;
  Matches   89;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
             |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db         1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKYLITKELPNTLKNTKDQGTIDKIFYNL 60

Qy        61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
             |||||||||||||||||||||||||||||||
Db        61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
```

```
RESULT 8
US-09-646-651B-1
, Sequence 1, Application US/09646651B
, GENERAL INFORMATION:
,  APPLICANT: Kiesewetter, Stefan
,  APPLICANT:  Kuhn, Eckehard
,  APPLICANT:  Koch-Pelster, Brigitte
,  APPLICANT:  Brunner, Herwig
,  TITLE OF INVENTION: METAL-CONTAINING RIBONUCLEOTIDE POLYPEPTIDES
;  FILE REFERENCE: 206579
;  CURRENT APPLICATION NUMBER: US/09/646,651B
;  CURRENT FILING DATE:  2000-09-13
;  PRIOR APPLICATION NUMBER: PCT/EP98/07722
;  PRIOR FILING DATE: 1998-11-30
;  PRIOR APPLICATION NUMBER: DE 198 11 047.2
;  PRIOR FILING DATE: 1998-03-13
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 91
;   TYPE: PRT
;   ORGANISM: Unknown
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: ()..()
;   OTHER INFORMATION: Angiotropin-related protein
US-09-646-651B-1

  Query Match             97.4%;  Score 455;  DB 20;  Length 91;
  Best Local Similarity   97.8%;  Pred. No. 4.1e-43;
  Matches   89;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||| ||
Db          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKYLITKELPNTLKNTKDQGTIDKIFYNL 60

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              |||||||||||||||||||||||||||||||
Db         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91


RESULT 9
US-09-791-537-99618
; Sequence 99618, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 99618
;   LENGTH: 92
;   TYPE: PRT
;   ORGANISM: Bos taurus
US-09-791-537-99618

  Query Match             79.4%;  Score 371;  DB 22;  Length 92;
  Best Local Similarity   81.3%;  Pred. No. 1.3e-33;
  Matches   74;  Conservative    9;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              |||||||||||||||||||:||:||| |||||||||||| ||:||||| |||||||:|
Db          2 TKLEDHLEGIINIFHQYSVRVGHFDTLNKRELKQLITKELPKTLQNTKDQPTIDKIFQDL 61

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              ||::|  |||:||||||: || ||| :||||
Db         62 DADKDGAVSFEEFVVLVSRVLKTAHIDIHKE 92


RESULT 10
US-09-167-705-3
; Sequence 3, Application US/09167705B
; GENERAL INFORMATION:
;  APPLICANT: Schmidt, Ann Marie
;  APPLICANT: Stern, David
;  TITLE OF INVENTION: Extracellular Novel RAGE Binding Protein (EN-RAGE) and
;  TITLE OF INVENTION: Uses Thereof
;  FILE REFERENCE: 0575/55873
;  CURRENT APPLICATION NUMBER: US/09/167,705B
;  CURRENT FILING DATE: 1998-10-06
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 90
;   TYPE: PRT
;   ORGANISM: Human
US-09-167-705-3

  Query Match             78.4%;  Score 366;  DB 15;  Length 90;
  Best Local Similarity   81.1%;  Pred. No. 4.7e-33;
  Matches   73;  Conservative    9;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              |||||||||||||||||||:||:||| |||||||||||| ||:||||| |||||||:|
Db          1 TKLEDHLEGIINIFHQYSVRVGHFDTLNKRELKQLITKELPKTLQNTKDQPTIDKIFQDL 60

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
              ||::|  |||:||||||: || ||| :|||
Db         61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90


RESULT 11
US-09-872-185-11
; Sequence 11, Application US/09872185
; GENERAL INFORMATION:
;  APPLICANT: Stern, David M.
;  APPLICANT:  Herold, Kevan
;  APPLICANT:  Yan, Shi Du
;  APPLICANT:  Schmidt, Ann Marie
;  APPLICANT:  Lamster, Ira
;  TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION
;  FILE REFERENCE: 0575/64080
;  CURRENT APPLICATION NUMBER: US/09/872,185
;  CURRENT FILING DATE:  2001-08-20
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
;   LENGTH: 90
;   TYPE: PRT
;   ORGANISM: Bovine
US-09-872-185-11

  Query Match             78.4%;  Score 366;  DB 23;  Length 90;
  Best Local Similarity   81.1%;  Pred. No. 4.7e-33;
  Matches   73;  Conservative    9;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              |||||||||||||||||||:||:||| |||||||||||| ||:||||| |||||||:|
Db          1 TKLEDHLEGIINIFHQYSVRVGHFDTLNKRELKQLITKELPKTLQNTKDQPTIDKIFQDL 60

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
              ||::|  |||:||||||: || ||| :|||
Db         61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90


RESULT 12
US-09-872-185-12
; Sequence 12, Application US/09872185
; GENERAL INFORMATION:
;  APPLICANT: Stern, David M.
;  APPLICANT:  Herold, Kevan
```

```
  APPLICANT:  Yan, Shi Du
  APPLICANT:  Schmidt, Ann Marie
  APPLICANT:  Lamster, Ira
  TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION
  FILE REFERENCE: 0575/64080
  CURRENT APPLICATION NUMBER: US/09/872,185
  CURRENT FILING DATE:  2001-08-20
  NUMBER OF SEQ ID NOS: 16
  SOFTWARE: PatentIn version 3.1
 SEQ ID NO 12
   LENGTH: 90
   TYPE: PRT
   ORGANISM: Bovine
S-09-872-185-12

 Query Match             78.4%;  Score 366;  DB 23;  Length 90;
 Best Local Similarity   81.1%;  Pred. No. 4.7e-33;
 Matches   73;  Conservative    9;  Mismatches    8;  Indels    0;  Gaps    0;

y           1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              |||||||||||||||||||||:||:||| |||||||||||| ||:||||| |||||||:|
b           1 TKLEDHLEGIINIFHQYSVRVGHFDTLNKRELKQLITKELPKTLQNTKDQPTIDKIFQDL 60

y          61 DANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
              ||::|  |||:||||||: || ||| :|||
b          61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90


ESULT 13
S-09-872-185B-11
 Sequence 11, Application US/09872185B
 GENERAL INFORMATION:
  APPLICANT: Stern, David M.
  APPLICANT:  Herold, Kevan
  APPLICANT:  Yan, Shi Du
  APPLICANT:  Schmidt, Ann Marie
  APPLICANT:  Lamster, Ira
  TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION
  FILE REFERENCE: 0575/64080
  CURRENT APPLICATION NUMBER: US/09/872,185B
  CURRENT FILING DATE:  2001-06-01
  NUMBER OF SEQ ID NOS: 16
  SOFTWARE: PatentIn version 3.1
 SEQ ID NO 11
   LENGTH: 90
   TYPE: PRT
   ORGANISM: Bovine
S-09-872-185B-11

 Query Match             78.4%;  Score 366;  DB 23;  Length 90;
 Best Local Similarity   81.1%;  Pred. No. 4.7e-33;
 Matches   73;  Conservative    9;  Mismatches    8;  Indels    0;  Gaps    0;

y           1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              |||||||||||||||||||||:||:||| |||||||||||| ||:||||| |||||||:|
b           1 TKLEDHLEGIINIFHQYSVRVGHFDTLNKRELKQLITKELPKTLQNTKDQPTIDKIFQDL 60

y          61 DANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
              ||::|  |||:||||||: || ||| :|||
b          61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90


ESULT 14
S-09-872-185B-12
 Sequence 12, Application US/09872185B
 GENERAL INFORMATION:
  APPLICANT: Stern, David M.
  APPLICANT:  Herold, Kevan
  APPLICANT:  Yan, Shi Du
  APPLICANT:  Schmidt, Ann Marie
  APPLICANT:  Lamster, Ira
  TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION
  FILE REFERENCE: 0575/64080
  CURRENT APPLICATION NUMBER: US/09/872,185B
  CURRENT FILING DATE:  2001-06-01
  NUMBER OF SEQ ID NOS: 16
  SOFTWARE: PatentIn version 3.1
 SEQ ID NO 12
   LENGTH: 90
   TYPE: PRT
   ORGANISM: Bovine
US-09-872-185B-12

  Query Match             78.4%;  Score 366;  DB 23;  Length 90;
  Best Local Similarity   81.1%;  Pred. No. 4.7e-33;
  Matches   73;  Conservative    9;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              |||||||||||||||||||||:||:||| |||||||||||| ||:||||| |||||||:|
Db          1 TKLEDHLEGIINIFHQYSVRVGHFDTLNKRELKQLITKELPKTLQNTKDQPTIDKIFQDL 60

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
              ||::|  |||:||||||: || ||| :|||
Db         61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90


RESULT 15
US-10-666-513-3
; Sequence 3, Application US/10666513
; GENERAL INFORMATION:
;  APPLICANT: Schmidt, Ann Marie
;  APPLICANT:  Stern, David
;  TITLE OF INVENTION: Extracellular Novel RAGE Binding Protein (EN-RAGE) and
;  TITLE OF INVENTION:  Uses Thereof
;  FILE REFERENCE: 0575/55873
;  CURRENT APPLICATION NUMBER: US/10/666,513
;  CURRENT FILING DATE:  2003-09-19
;  PRIOR APPLICATION NUMBER: US/09/167,705B
;  PRIOR FILING DATE: 1998-10-06
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 90
;   TYPE: PRT
;   ORGANISM: Human
US-10-666-513-3

  Query Match             78.4%;  Score 366;  DB 31;  Length 90;
  Best Local Similarity   81.1%;  Pred. No. 4.7e-33;
  Matches   73;  Conservative    9;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              |||||||||||||||||||||:||:||| |||||||||||| ||:||||| |||||||:|
Db          1 TKLEDHLEGIINIFHQYSVRVGHFDTLNKRELKQLITKELPKTLQNTKDQPTIDKIFQDL 60

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
              ||::|  |||:||||||: || ||| :|||
Db         61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90


Search completed: March  5, 2004, 10:39:08
Job time : 183 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

M protein - protein search, using sw model

un on: March 5, 2004, 10:31:35 ; Search time 10 Seconds
(without alignments)
91.517 Million cell updates/sec

itle: US-09-646-651C-1
erfect score: 467
equence: 1 TKLEDHLEGIINIFHQYSVR.....EFVVLVTDVLITAHDNIHKE 91

coring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

earched: 105621 seqs, 10056811 residues

otal number of hits satisfying chosen parameters: 105621

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

. Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

esult No.	Score	% Query Match	Length	DB	ID	Description
1	366	78.4	90	6	US-10-665-867-3	Sequence 3, Appli
2	366	78.4	90	6	US-10-665-867-4	Sequence 4, Appli
3	206	44.1	50	6	US-10-665-867-2	Sequence 2, Appli
4	159	34.0	89	6	US-10-624-631-32	Sequence 32, Appl
5	151	32.3	93	6	US-10-624-631-31	Sequence 31, Appl
6	146	31.3	98	6	US-10-624-631-28	Sequence 28, Appl
7	134.5	28.8	95	6	US-10-487-337-4	Sequence 4, Appli
8	134	28.7	90	6	US-10-624-631-29	Sequence 29, Appl
9	133	28.5	99	6	US-10-624-631-33	Sequence 33, Appl
10	129.5	27.7	97	6	US-10-487-337-8	Sequence 8, Appli
11	127	27.2	88	6	US-10-624-631-34	Sequence 34, Appl
12	124	26.6	495	6	US-10-767-471-717	Sequence 717, App
13	109	23.3	61	6	US-10-624-631-30	Sequence 30, Appl
14	102	21.8	104	1	PCT-US04-02188-178	Sequence 178, App
15	102	21.8	104	6	US-10-764-425-178	Sequence 178, App
16	78	16.7	92	7	US-60-543-011-10	Sequence 10, Appl
17	74.5	16.0	135	6	US-10-767-701-38583	Sequence 38583, A
18	74.5	16.0	223	6	US-10-767-701-40040	Sequence 40040, A
19	72	15.4	532	6	US-10-451-467A-312	Sequence 312, App
20	71	15.2	226	6	US-10-767-701-39009	Sequence 39009, A
21	70.5	15.1	163	6	US-10-767-701-36974	Sequence 36974, A
22	66.5	14.2	305	6	US-10-767-701-44602	Sequence 44602, A
23	66	14.1	629	6	US-10-667-290-2	Sequence 2, Appli
24	66	14.1	629	6	US-10-667-289-2	Sequence 2, Appli
25	65.5	14.0	498	6	US-10-767-701-46009	Sequence 46009, A
26	65	13.9	95	6	US-10-767-701-61761	Sequence 61761, A
27	64.5	13.8	198	6	US-10-767-701-61992	Sequence 61992, A
28	64.5	13.8	386	6	US-10-417-884A-5555	Sequence 5555, Ap
29	64	13.7	167	6	US-10-767-701-40255	Sequence 40255, A
30	63.5	13.6	155	6	US-10-767-701-49565	Sequence 49565, A
31	63.5	13.6	244	5	US-09-830-230A-506	Sequence 506, App
32	63.5	13.6	264	5	US-09-830-230A-505	Sequence 505, App
33	62	13.3	623	6	US-10-416-330-38	Sequence 38, Appl
34	61.5	13.2	330	6	US-10-767-701-38434	Sequence 38434, A
35	61.5	13.2	395	6	US-10-451-467A-224	Sequence 224, App
36	61	13.1	561	6	US-10-417-884A-4239	Sequence 4239, Ap
37	61	13.1	873	1	PCT-US04-03291-37	Sequence 37, Appl
38	61	13.1	873	6	US-10-771-931-37	Sequence 37, Appl
39	60.5	13.0	319	6	US-10-417-884A-4910	Sequence 4910, Ap
40	60.5	13.0	319	6	US-10-417-884A-4912	Sequence 4912, Ap
41	60	12.8	780	6	US-10-417-884A-5050	Sequence 5050, Ap
42	59.5	12.7	179	6	US-10-767-701-57231	Sequence 57231, A
43	59.5	12.7	242	6	US-10-417-884A-4507	Sequence 4507, Ap
44	59.5	12.7	242	6	US-10-417-884A-4508	Sequence 4508, Ap
45	59.5	12.7	255	6	US-10-417-884A-3995	Sequence 3995, Ap

ALIGNMENTS

```
RESULT 1
US-10-665-867-3
; Sequence 3, Application US/10665867
; GENERAL INFORMATION:
;  APPLICANT: Schmidt, Ann Marie
;  APPLICANT:  Stern, David
;  TITLE OF INVENTION: EXTRACELLULAR NOVEL RAGE BINDING PROTEIN (EN-RAGE) AND USES THERE
;  FILE REFERENCE: 0575/55873-B-PCT-US
;  CURRENT APPLICATION NUMBER: US/10/665,867
;  CURRENT FILING DATE:  2003-09-17
;  PRIOR APPLICATION NUMBER: US/09/826,589
;  PRIOR FILING DATE: 2001-04-05
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 90
;   TYPE: PRT
;   ORGANISM: Bovine
US-10-665-867-3

  Query Match             78.4%;  Score 366;  DB 6;  Length 90;
  Best Local Similarity   81.1%;  Pred. No. 3.6e-34;
  Matches   73;  Conservative    9;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              |||||||||||||||||||:||:||| |||||||||||| ||:||||| |||||||:|
Db          1 TKLEDHLEGIINIFHQYSVRVGHFDTLNKRELKQLITKELPKTLQNTKDQPTIDKIFQDL 60

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
                ||::|  |||:|||||: || ||| :|||
Db         61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90


RESULT 2
US-10-665-867-4
; Sequence 4, Application US/10665867
; GENERAL INFORMATION:
;  APPLICANT: Schmidt, Ann Marie
;  APPLICANT:  Stern, David
;  TITLE OF INVENTION: EXTRACELLULAR NOVEL RAGE BINDING PROTEIN (EN-RAGE) AND USES THERE
;  FILE REFERENCE: 0575/55873-B-PCT-US
;  CURRENT APPLICATION NUMBER: US/10/665,867
;  CURRENT FILING DATE:  2003-09-17
;  PRIOR APPLICATION NUMBER: US/09/826,589
;  PRIOR FILING DATE: 2001-04-05
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 90
```

```
;   TYPE: PRT
;   ORGANISM: Bovine
US-10-665-867-4

  Query Match             78.4%;  Score 366;  DB 6;  Length 90;
  Best Local Similarity   81.1%;  Pred. No. 3.6e-34;
  Matches   73;  Conservative    9;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              |||||||||||||||||||:||:||| |||||||||||| ||:|||| |||||||:|
Db          1 TKLEDHLEGIINIFHQYSVRVGHFDTLNKRELKQLITKELPKTLQNTKDQPTIDKIFQDL 60

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
              ||::|  |||:||||||: || ||| :|||
Db         61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90


RESULT 3
US-10-665-867-2
; Sequence 2, Application US/10665867
; GENERAL INFORMATION:
;  APPLICANT: Schmidt, Ann Marie
;  APPLICANT:  Stern, David
;  TITLE OF INVENTION: EXTRACELLULAR NOVEL RAGE BINDING PROTEIN (EN-RAGE) AND USES THERE
;  FILE REFERENCE: 0575/55873-B-PCT-US
;  CURRENT APPLICATION NUMBER: US/10/665,867
;  CURRENT FILING DATE:  2003-09-17
;  PRIOR APPLICATION NUMBER: US/09/826,589
;  PRIOR FILING DATE: 2001-04-05
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 50
;   TYPE: PRT
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (47)..(47)
;   OTHER INFORMATION: x=any amino acid
US-10-665-867-2

  Query Match             44.1%;  Score 206;  DB 6;  Length 50;
  Best Local Similarity   82.0%;  Pred. No. 1e-16;
  Matches   41;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQ 50
              |||||||||||| |||||:||:||| | ||||| ||||| ||:| |||
Db          1 TKLEDHLEGIINIGHQYSVRVGHFDTLNKYELKQLGTKELPKTLQNXKDQ 50


RESULT 4
US-10-624-631-32
; Sequence 32, Application US/10624631
; GENERAL INFORMATION:
;  APPLICANT: Morphotek, Inc.
;  APPLICANT:  Grasso, Luigi
;  APPLICANT:  Kline, J. Bradford
;  APPLICANT:  Nicolaides, Nicholas C.
;  APPLICANT:  Sass, Philip M.
;  TITLE OF INVENTION: Methods for Generating Enhanced Antibody-Producing Cell Lines wit
;  TITLE OF INVENTION:  Improved Growth Characteristics
;  FILE REFERENCE: MOR-0241/HD0002 US
;  CURRENT APPLICATION NUMBER: US/10/624,631
;  CURRENT FILING DATE:  2003-07-21
;  PRIOR APPLICATION NUMBER: 60/397,027
;  PRIOR FILING DATE: 2002-07-19
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
;   LENGTH: 89
;   TYPE: PRT
;   ORGANISM: Rattus norvegicus
US-10-624-631-32

  Query Match             34.0%;  Score 159;  DB 6;  Length 89;
  Best Local Similarity   33.0%;  Pred. No. 3.9e-11;
  Matches   30;  Conservative   23;  Mismatches   34;  Indels    4;  Gaps    1;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              |:|| |  :| ::| ||   |::  | : : ::::| | |  ::|       : :|: |
Db          3 TELEKALSNVIEVYHNYSGIKGNHHALYRDDFRKMVTTECPQFVQNK----NTESLFKEL 58

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              | | |  ::|:||: ||  | : || : |||
Db         59 DVNSDNAINFEEFLALVIRVGVAAHKDSHKE 89


RESULT 5
US-10-624-631-31
; Sequence 31, Application US/10624631
; GENERAL INFORMATION:
;  APPLICANT: Morphotek, Inc.
;  APPLICANT:  Grasso, Luigi
;  APPLICANT:  Kline, J. Bradford
;  APPLICANT:  Nicolaides, Nicholas C.
;  APPLICANT:  Sass, Philip M.
;  TITLE OF INVENTION: Methods for Generating Enhanced Antibody-Producing Cell Lines wit
;  TITLE OF INVENTION:  Improved Growth Characteristics
;  FILE REFERENCE: MOR-0241/HD0002 US
;  CURRENT APPLICATION NUMBER: US/10/624,631
;  CURRENT FILING DATE:  2003-07-21
;  PRIOR APPLICATION NUMBER: 60/397,027
;  PRIOR FILING DATE: 2002-07-19
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
;   LENGTH: 93
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-624-631-31

  Query Match             32.3%;  Score 151;  DB 6;  Length 93;
  Best Local Similarity   35.2%;  Pred. No. 3.2e-10;
  Matches   32;  Conservative   24;  Mismatches   31;  Indels    4;  Gaps    2;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              |:|| |  ||:::|:||:  |::  : : :||:|:  | |  ::    :|  |  |: |
Db          3 TELEKALNSIIDVYHKYSLIKGNFHAVYRDDLKKLLETECPQYIRK---KGA-DVWFKEL 58

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              | | |  |:|:||::||  : : ||   |:|
Db         59 DINTDGAVNFQEFLILVIKMGVAAHKKSHEE 89


RESULT 6
US-10-624-631-28
; Sequence 28, Application US/10624631
; GENERAL INFORMATION:
;  APPLICANT: Morphotek, Inc.
;  APPLICANT:  Grasso, Luigi
;  APPLICANT:  Kline, J. Bradford
;  APPLICANT:  Nicolaides, Nicholas C.
;  APPLICANT:  Sass, Philip M.
;  TITLE OF INVENTION: Methods for Generating Enhanced Antibody-Producing Cell Lines wit
;  TITLE OF INVENTION:  Improved Growth Characteristics
;  FILE REFERENCE: MOR-0241/HD0002 US
;  CURRENT APPLICATION NUMBER: US/10/624,631
;  CURRENT FILING DATE:  2003-07-21
;  PRIOR APPLICATION NUMBER: 60/397,027
;  PRIOR FILING DATE: 2002-07-19
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
;   LENGTH: 98
;   TYPE: PRT
;   ORGANISM: Mus musculus
```

/S-10-624-631-28

Query Match 31.3%; Score 146; DB 6; Length 98;
Best Local Similarity 35.6%; Pred. No. 1.3e-09;
Matches 31; Conservative 18; Mismatches 38; Indels 0; Gaps 0;

```
)y          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              |: |  :| :| :| :|| : |:   | | |    :  ||    || || | :|:: : |
)b          3 TETERCIESLIAVFQKYSGKDGNNTQLSKTEFLSFMNTELAAFTKNQKDPGVLDRMMKKL 62

)y         61 DANQDEQVSFKEFVVLVTDVLITAHDN 87
              | | | |: |:||: |:  : |  ||:
)b         63 DLNCDGQLDFQEFLNLIGGLAIACHDS 89
```

!ESULT 7
/S-10-487-337-4
Sequence 4, Application US/10487337
GENERAL INFORMATION:
APPLICANT: Kenji OKUSE
APPLICANT: Mark BAKER
APPLICANT: Louis POON
APPLICANT: John Nicholas WOOD
APPLICANT: Misbah MALIK-HALL
TITLE OF INVENTION: SODIUM CHANNEL REGULATORS AND MODULATORS
FILE REFERENCE: 117-492 / N.86242A GCW
CURRENT APPLICATION NUMBER: US/10/487,337
CURRENT FILING DATE: 2004-02-20
PRIOR APPLICATION NUMBER: PCT/GB02/03852
PRIOR FILING DATE: 2002-08-20
PRIOR APPLICATION NUMBER: GB 0120238.1
PRIOR FILING DATE: 2001-08-20
NUMBER OF SEQ ID NOS: 55
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 95
TYPE: PRT
ORGANISM: Rattus norvegicus
/S-10-487-337-4

Query Match 28.8%; Score 134.5; DB 6; Length 95;
Best Local Similarity 32.6%; Pred. No. 2.3e-08;
Matches 28; Conservative 25; Mismatches 30; Indels 3; Gaps 1;

```
)y          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              :::|  :| ::  ||:::   |  : | | :|: |: :| |  |:| ||   :||| ::|
)b          3 SQMEHAMETMMLTFHRFA---GEKNYLTKEDLRVLMEREFPGFLENQKDPLAVDKIMKDL 59

)y         61 DANQDEQVSFKEFVVLVTDVLITAHD 86
              |  :| :| |: |: ||  ::|  :|
)b         60 DQCRDGKVGFQSFLSLVAGLIIACND 85
```

!ESULT 8
/S-10-624-631-29
Sequence 29, Application US/10624631
GENERAL INFORMATION:
APPLICANT: Morphotek, Inc.
APPLICANT: Grasso, Luigi
APPLICANT: Kline, J. Bradford
APPLICANT: Nicolaides, Nicholas C.
APPLICANT: Sass, Philip M.
TITLE OF INVENTION: Methods for Generating Enhanced Antibody-Producing Cell Lines wit
TITLE OF INVENTION: Improved Growth Characteristics
FILE REFERENCE: MOR-0241/HD0002 US
CURRENT APPLICATION NUMBER: US/10/624,631
CURRENT FILING DATE: 2003-07-21
PRIOR APPLICATION NUMBER: 60/397,027
PRIOR FILING DATE: 2002-07-19
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn version 3.1
SEQ ID NO 29
LENGTH: 90
; TYPE: PRT
; ORGANISM: Canis familiaris
US-10-624-631-29

Query Match 28.7%; Score 134; DB 6; Length 90;
Best Local Similarity 34.6%; Pred. No. 2.4e-08;
Matches 27; Conservative 18; Mismatches 33; Indels 0; Gaps 0;

```
Qy         10 IINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNLDANQDEQVS 69
              :| :| ::: : |:  || | |    :  ||    || || | :|:: : || | | |:
Db          4 LIAVFQKFAGKEGNNCTLSKTEFLTFMNTELAAFTKNQKDPGVLDRMMKKLDLNSDGQLD 63

Qy         70 FKEFVVLVTDVLITAHDN 87
              |:||: |:  : |  ||:
Db         64 FQEFLNLIGGMAIACHDS 81
```

RESULT 9
US-10-624-631-33
; Sequence 33, Application US/10624631
; GENERAL INFORMATION:
; APPLICANT: Morphotek, Inc.
; APPLICANT: Grasso, Luigi
; APPLICANT: Kline, J. Bradford
; APPLICANT: Nicolaides, Nicholas C.
; APPLICANT: Sass, Philip M.
; TITLE OF INVENTION: Methods for Generating Enhanced Antibody-Producing Cell Lines wit
; TITLE OF INVENTION: Improved Growth Characteristics
; FILE REFERENCE: MOR-0241/HD0002 US
; CURRENT APPLICATION NUMBER: US/10/624,631
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: 60/397,027
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Sus scrofa
US-10-624-631-33

Query Match 28.5%; Score 133; DB 6; Length 99;
Best Local Similarity 32.2%; Pred. No. 3.6e-08;
Matches 29; Conservative 21; Mismatches 40; Indels 0; Gaps 0;

```
Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              |: |  :| :| || ::: | |:   :  | |    :  ||    :| || | :|:: : |
Db          6 TETERCIESLIAIFQKHAGRDGNNTKISKTEFLIFMNTELAAFTQNQKDPGVLDRMMKKL 65

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
              | : | |: |:||: |:  : |  ||:  |
Db         66 DLDSDGQLDFQEFLNLIGGLAIACHDSFIK 95
```

RESULT 10
US-10-487-337-8
; Sequence 8, Application US/10487337
; GENERAL INFORMATION:
; APPLICANT: Kenji OKUSE
; APPLICANT: Mark BAKER
; APPLICANT: Louis POON
; APPLICANT: John Nicholas WOOD
; APPLICANT: Misbah MALIK-HALL
; TITLE OF INVENTION: SODIUM CHANNEL REGULATORS AND MODULATORS
; FILE REFERENCE: 117-492 / N.86242A GCW
; CURRENT APPLICATION NUMBER: US/10/487,337
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: PCT/GB02/03852
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: GB 0120238.1
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 8
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-487-337-8

Query Match 27.7%; Score 129.5; DB 6; Length 97;
Best Local Similarity 32.6%; Pred. No. 8.5e-08;
Matches 28; Conservative 23; Mismatches 32; Indels 3; Gaps 1;

Qy 1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
Db 3 SQMEHAMETMMFTFHKFAGDKGY---LTKEDLRVLMEKEFPGFLENQKDPLAVDKIMKDL 59

Qy 61 DANQDEQVSFKEFVVLVTDVLITAHD 86
Db 60 DQCRDGKVGFQSFFSLIAGLTIACND 85

RESULT 11
US-10-624-631-34
; Sequence 34, Application US/10624631
; GENERAL INFORMATION:
; APPLICANT: Morphotek, Inc.
; APPLICANT: Grasso, Luigi
; APPLICANT: Kline, J. Bradford
; APPLICANT: Nicolaides, Nicholas C.
; APPLICANT: Sass, Philip M.
; TITLE OF INVENTION: Methods for Generating Enhanced Antibody-Producing Cell Lines wit
; TITLE OF INVENTION: Improved Growth Characteristics
; FILE REFERENCE: MOR-0241/HD0002 US
; CURRENT APPLICATION NUMBER: US/10/624,631
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: 60/397,027
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Consensus Sequence
US-10-624-631-34

Query Match 27.2%; Score 127; DB 6; Length 88;
Best Local Similarity 34.6%; Pred. No. 1.4e-07;
Matches 27; Conservative 17; Mismatches 32; Indels 2; Gaps 1;

Qy 10 IINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNLDANQDEQVS 69
Db 8 LIAVFQKYAGKDG--NNLSKTEFLSFMNTELAAFTKNQKDPGVLDRMMKKLDLNSDGQLD 65

Qy 70 FKEFVVLVTDVLITAHDN 87
Db 66 FQEFLNLIGGLAIACHDS 83

RESULT 12
US-10-767-471-717
; Sequence 717, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 717
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-767-471-717

Query Match 26.6%; Score 124; DB 6; Length 495;
Best Local Similarity 34.2%; Pred. No. 3.5e-06;
Matches 27; Conservative 18; Mismatches 34; Indels 0; Gaps 0;

Qy 2 KLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNLD 61
Db 3 QLLQNINGIIEAFRRYARTEGNCTALTRGELKRLLEQEFADVIVKPHDPATVDEVLRLLD 62

Qy 62 ANQDEQVSFKEFVVLVTDV 80
Db 63 EDHTGTVEFKEFLVLVFKV 81

RESULT 13
US-10-624-631-30
; Sequence 30, Application US/10624631
; GENERAL INFORMATION:
; APPLICANT: Morphotek, Inc.
; APPLICANT: Grasso, Luigi
; APPLICANT: Kline, J. Bradford
; APPLICANT: Nicolaides, Nicholas C.
; APPLICANT: Sass, Philip M.
; TITLE OF INVENTION: Methods for Generating Enhanced Antibody-Producing Cell Lines wit
; TITLE OF INVENTION: Improved Growth Characteristics
; FILE REFERENCE: MOR-0241/HD0002 US
; CURRENT APPLICATION NUMBER: US/10/624,631
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: 60/397,027
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-624-631-30

Query Match 23.3%; Score 109; DB 6; Length 61;
Best Local Similarity 37.3%; Pred. No. 8.4e-06;
Matches 22; Conservative 11; Mismatches 26; Indels 0; Gaps 0;

Qy 13 IFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNLDANQDEQVSFK 71
Db 3 VFQKYAGKDGHSVTLSKTEFLSFMNTELAAFTKNQKDPGVLDRMMKKLDLNSDGQLDFQ 61

RESULT 14
PCT-US04-02188-178
; Sequence 178, Application PC/TUS0402188
; GENERAL INFORMATION:
; APPLICANT: Bayer Pharmaceuticals Corporation
; APPLICANT: Eveleigh, Deepa
; APPLICANT: Bigwood, Douglas
; APPLICANT: Taylor, Ian
; TITLE OF INVENTION: EXPRESSION PROFILES FOR COLON CANCER AND METHODS OF USE
; FILE REFERENCE: 5151
; CURRENT APPLICATION NUMBER: PCT/US04/02188
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: 60/442,582
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 178
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US04-02188-178

Query Match 21.8%; Score 102; DB 1; Length 104;
Best Local Similarity 36.0%; Pred. No. 0.00011;

Matches 27; Conservative 18; Mismatches 26; Indels 4; Gaps 3;

```
)y          3 LEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNLDA 62
              :|  :| :|  ||||||  |  :||   ||: |:|::||: : :  : |  :||  || :
)b         18 VERAIETLIKNFHQYSVE-GGKETLTPSELRDLVTQQLPHLMPS--NCGLEEKI-ANLGS 73

)y         63 NQDEQVSFKEFVVLV 77
                | :: |: |  |:
)b         74 CNDSKLEFRSFWELI 88
```

!ESULT 15
!S-10-764-425-178
Sequence 178, Application US/10764425
GENERAL INFORMATION:
APPLICANT: Bayer Pharmaceuticals Corporation
APPLICANT: Eveleigh, Deepa
APPLICANT: Bigwood, Douglas
APPLICANT: Taylor, Ian
TITLE OF INVENTION: EXPRESSION PROFILES FOR COLON CANCER AND METHODS OF USE
FILE REFERENCE: 5151
CURRENT APPLICATION NUMBER: US/10/764,425
CURRENT FILING DATE: 2004-01-23
PRIOR APPLICATION NUMBER: 60/442,582
PRIOR FILING DATE: 2003-01-24
NUMBER OF SEQ ID NOS: 191
SOFTWARE: PatentIn version 3.2
SEQ ID NO 178
LENGTH: 104
TYPE: PRT
ORGANISM: Homo sapiens
!S-10-764-425-178

Query Match 21.8%; Score 102; DB 6; Length 104;
Best Local Similarity 36.0%; Pred. No. 0.00011;
Matches 27; Conservative 18; Mismatches 26; Indels 4; Gaps 3;

```
y           3 LEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNLDA 62
              :|  :| :|  ||||||  |  :||   ||: |:|::||: : :  : |  :||  || :
b          18 VERAIETLIKNFHQYSVE-GGKETLTPSELRDLVTQQLPHLMPS--NCGLEEKI-ANLGS 73

y          63 NQDEQVSFKEFVVLV 77
                | :: |: |  |:
b          74 CNDSKLEFRSFWELI 88
```

earch completed: March 5, 2004, 10:39:30
ob time : 10 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on: March 8, 2004, 04:45:44 ; Search time 3135 Seconds
(without alignments)
1200.930 Million cell updates/sec

Title: US-09-646-651C-2
Perfect score: 107
Sequence: 1 ggaaaaunnnnnunauaugn.....nnnnnnnnnnnnnnuuagcag 107

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 37577330 seqs, 17593059518 residues

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main:*
1: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US080_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US081_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US082_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US083_COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US084_COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US085_COMB.seq:*
10: /cgn2_6/ptodata/2/pna/US086_COMB.seq:*
11: /cgn2_6/ptodata/2/pna/US087_COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US088_COMB.seq:*
13: /cgn2_6/ptodata/2/pna/US089_COMB.seq:*
14: /cgn2_6/ptodata/2/pna/US090_COMB.seq:*
15: /cgn2_6/ptodata/2/pna/US091_COMB.seq:*
16: /cgn2_6/ptodata/2/pna/US092A_COMB.seq:*
17: /cgn2_6/ptodata/2/pna/US092B_COMB.seq:*
18: /cgn2_6/ptodata/2/pna/US093A_COMB.seq:*
19: /cgn2_6/ptodata/2/pna/US093B_COMB.seq:*
20: /cgn2_6/ptodata/2/pna/US094_COMB.seq:*
21: /cgn2_6/ptodata/2/pna/US095A_COMB.seq:*
22: /cgn2_6/ptodata/2/pna/US095B_COMB.seq:*
23: /cgn2_6/ptodata/2/pna/US095C_COMB.seq:*
24: /cgn2_6/ptodata/2/pna/US095D_COMB.seq:*
25: /cgn2_6/ptodata/2/pna/US096A_COMB.seq:*
26: /cgn2_6/ptodata/2/pna/US096B_COMB.seq:*
27: /cgn2_6/ptodata/2/pna/US096C_COMB.seq:*
28: /cgn2_6/ptodata/2/pna/US096D_COMB.seq:*
29: /cgn2_6/ptodata/2/pna/US096E_COMB.seq:*
30: /cgn2_6/ptodata/2/pna/US097A_COMB.seq:*
31: /cgn2_6/ptodata/2/pna/US097B_COMB.seq:*
32: /cgn2_6/ptodata/2/pna/US097C_COMB.seq:*
33: /cgn2_6/ptodata/2/pna/US098A_COMB.seq:*
34: /cgn2_6/ptodata/2/pna/US098B_COMB.seq:*
35: /cgn2_6/ptodata/2/pna/US098C_COMB.seq:*
36: /cgn2_6/ptodata/2/pna/US098D_COMB.seq:*
37: /cgn2_6/ptodata/2/pna/US099A_COMB.seq:*
38: /cgn2_6/ptodata/2/pna/US099B_COMB.seq:*
39: /cgn2_6/ptodata/2/pna/US099C_COMB.seq:*
40: /cgn2_6/ptodata/2/pna/US099D_COMB.seq:*
41: /cgn2_6/ptodata/2/pna/US099E_COMB.seq:*
42: /cgn2_6/ptodata/2/pna/US099F_COMB.seq:*
43: /cgn2_6/ptodata/2/pna/US100A_COMB.seq:*
44: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
45: /cgn2_6/ptodata/2/pna/US101A_COMB.seq:*
46: /cgn2_6/ptodata/2/pna/US101B_COMB.seq:*
47: /cgn2_6/ptodata/2/pna/US102A_COMB.seq:*
48: /cgn2_6/ptodata/2/pna/US102B_COMB.seq:*
49: /cgn2_6/ptodata/2/pna/US103A_COMB.seq:*
50: /cgn2_6/ptodata/2/pna/US103B_COMB.seq:*
51: /cgn2_6/ptodata/2/pna/US104A_COMB.seq:*
52: /cgn2_6/ptodata/2/pna/US104B_COMB.seq:*
53: /cgn2_6/ptodata/2/pna/US106_COMB.seq:*
54: /cgn2_6/ptodata/2/pna/US107A_COMB.seq:*
55: /cgn2_6/ptodata/2/pna/US107B_COMB.seq:*
56: /cgn2_6/ptodata/2/pna/US6000_COMB.seq:*
57: /cgn2_6/ptodata/2/pna/US6001_COMB.seq:*
58: /cgn2_6/ptodata/2/pna/US6002_COMB.seq:*
59: /cgn2_6/ptodata/2/pna/US6003_COMB.seq:*
60: /cgn2_6/ptodata/2/pna/US6004_COMB.seq:*
61: /cgn2_6/ptodata/2/pna/US6005_COMB.seq:*
62: /cgn2_6/ptodata/2/pna/US6006_COMB.seq:*
63: /cgn2_6/ptodata/2/pna/US6007_COMB.seq:*
64: /cgn2_6/ptodata/2/pna/US6008_COMB.seq:*
65: /cgn2_6/ptodata/2/pna/US6009_COMB.seq:*
66: /cgn2_6/ptodata/2/pna/US6010_COMB.seq:*
67: /cgn2_6/ptodata/2/pna/US6011_COMB.seq:*
68: /cgn2_6/ptodata/2/pna/US6012_COMB.seq:*
69: /cgn2_6/ptodata/2/pna/US6013_COMB.seq:*
70: /cgn2_6/ptodata/2/pna/US6014_COMB.seq:*
71: /cgn2_6/ptodata/2/pna/US6015_COMB.seq:*
72: /cgn2_6/ptodata/2/pna/US6016_COMB.seq:*
73: /cgn2_6/ptodata/2/pna/US6017_COMB.seq:*
74: /cgn2_6/ptodata/2/pna/US6018_COMB.seq:*
75: /cgn2_6/ptodata/2/pna/US6019_COMB.seq:*
76: /cgn2_6/ptodata/2/pna/US6020_COMB.seq:*
77: /cgn2_6/ptodata/2/pna/US6021_COMB.seq:*
78: /cgn2_6/ptodata/2/pna/US6022_COMB.seq:*
79: /cgn2_6/ptodata/2/pna/US6023A_COMB.seq:*
80: /cgn2_6/ptodata/2/pna/US6023B_COMB.seq:*
81: /cgn2_6/ptodata/2/pna/US6024_COMB.seq:*
82: /cgn2_6/ptodata/2/pna/US6025_COMB.seq:*
83: /cgn2_6/ptodata/2/pna/US6026_COMB.seq:*
84: /cgn2_6/ptodata/2/pna/US6027_COMB.seq:*
85: /cgn2_6/ptodata/2/pna/US6028_COMB.seq:*
86: /cgn2_6/ptodata/2/pna/US6029_COMB.seq:*
87: /cgn2_6/ptodata/2/pna/US6030_COMB.seq:*
88: /cgn2_6/ptodata/2/pna/US6031_COMB.seq:*
89: /cgn2_6/ptodata/2/pna/US6032_COMB.seq:*
90: /cgn2_6/ptodata/2/pna/US6033_COMB.seq:*
91: /cgn2_6/ptodata/2/pna/US6034_COMB.seq:*
92: /cgn2_6/ptodata/2/pna/US6035_COMB.seq:*
93: /cgn2_6/ptodata/2/pna/US6036_COMB.seq:*
94: /cgn2_6/ptodata/2/pna/US6037_COMB.seq:*
95: /cgn2_6/ptodata/2/pna/US6038_COMB.seq:*
96: /cgn2_6/ptodata/2/pna/US6039_COMB.seq:*
97: /cgn2_6/ptodata/2/pna/US6040_COMB.seq:*
98: /cgn2_6/ptodata/2/pna/US6041_COMB.seq:*
99: /cgn2_6/ptodata/2/pna/US6042_COMB.seq:*
100: /cgn2_6/ptodata/2/pna/US6043_COMB.seq:*
101: /cgn2_6/ptodata/2/pna/US6044_COMB.seq:*
102: /cgn2_6/ptodata/2/pna/US6045_COMB.seq:*
103: /cgn2_6/ptodata/2/pna/US6046_COMB.seq:*
104: /cgn2_6/ptodata/2/pna/US6047_COMB.seq:*
105: /cgn2_6/ptodata/2/pna/US6048_COMB.seq:*
106: /cgn2_6/ptodata/2/pna/US6049_COMB.seq:*
107: /cgn2_6/ptodata/2/pna/US6050_COMB.seq:*
108: /cgn2_6/ptodata/2/pna/US6051_COMB.seq:*
109: /cgn2_6/ptodata/2/pna/US6052_COMB.seq:*
110: /cgn2_6/ptodata/2/pna/US6053_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		49	45.8	107	27	US-09-646-651B-2	Sequence 2, Appli
2		49	45.8	107	27	US-09-646-651C-2	Sequence 2, Appli
3		31.4	29.3	298	27	US-09-646-651A-4	Sequence 4, Appli
4		31.4	29.3	298	27	US-09-646-651B-4	Sequence 4, Appli
5		31.4	29.3	298	27	US-09-646-651C-4	Sequence 4, Appli
6		29.8	27.9	226307	103	US-60-466-412-84605	Sequence 84605, A
7	C	29.4	27.5	1664976	1	PCT-US97-14900A-1	Sequence 1, Appli
8	C	29.4	27.5	1664976	29	US-09-692-570-1	Sequence 1, Appli
9	C	29.2	27.3	609	69	US-60-138-831-96	Sequence 96, Appl
10		29.2	27.3	903	70	US-60-140-806-338	Sequence 338, App
11		29.2	27.3	2334	70	US-60-141-856-255	Sequence 255, App
12		29.2	27.3	3020	73	US-60-173-464-20990	Sequence 20990, A
13		29.2	27.3	3020	75	US-60-191-637-25467	Sequence 25467, A
14		29.2	27.3	3020	75	US-60-191-681-20076	Sequence 20076, A
15		29.2	27.3	3238	70	US-60-142-844-270	Sequence 270, App
16		29.2	27.3	3456	25	US-09-614-150-25348	Sequence 25348, A
17		29.2	27.3	3456	25	US-09-614-150A-25348	Sequence 25348, A
18		29.2	27.3	4829	25	US-09-614-150-25336	Sequence 25336, A
19		29.2	27.3	4829	25	US-09-614-150A-25336	Sequence 25336, A
20		29.2	27.3	4829	73	US-60-173-464-20978	Sequence 20978, A
21		29.2	27.3	4829	75	US-60-191-637-25455	Sequence 25455, A
22		29.2	27.3	4829	75	US-60-191-681-20064	Sequence 20064, A
23		29.2	27.3	6062	72	US-60-167-217-9686	Sequence 9686, Ap
24	C	29.2	27.3	61423	21	US-09-528-237A-1870	Sequence 1870, Ap
25	C	29	27.1	27614	39	US-09-948-941-682	Sequence 682, App
26	C	29	27.1	35807	103	US-60-465-241-51556	Sequence 51556, A
27	C	28	26.2	183928	77	US-60-212-664-160	Sequence 160, App
28	C	27.8	26.0	510	25	US-09-605-702-3425	Sequence 3425, Ap
29	C	27.8	26.0	510	53	US-10-680-765-3425	Sequence 3425, Ap
30		27.4	25.6	474	104	US-60-475-872-2653	Sequence 2653, Ap
31		27.4	25.6	502	104	US-60-475-872-2591	Sequence 2591, Ap
32		27.4	25.6	524	1	PCT-US02-28214-280	Sequence 280, App
33		27.4	25.6	524	43	US-10-012-697-280	Sequence 280, App
34		27.4	25.6	524	84	US-60-275-688-280	Sequence 280, App
35		27.4	25.6	761	1	PCT-US02-28214-493	Sequence 493, App
36		27.4	25.6	761	43	US-10-012-697-493	Sequence 493, App
37		27.4	25.6	769	1	PCT-US02-28214-418	Sequence 418, App
38		27.4	25.6	769	43	US-10-012-697-418	GENERAL INFORMATI
39		27.4	25.6	769	84	US-60-275-688-419	Sequence 419, App
40		27.4	25.6	789	1	PCT-US02-28214-415	Sequence 415, App
41		27.4	25.6	789	43	US-10-012-697-415	Sequence 415, App
42		27.4	25.6	789	84	US-60-275-688-416	Sequence 416, App
43		27.4	25.6	835	84	US-60-275-688-494	Sequence 494, App
44		27.4	25.6	929	1	PCT-US02-28214-1307	Sequence 1307, Ap
45		27.4	25.6	929	43	US-10-012-697-1307	Sequence 1307, Ap

ALIGNMENTS

```
RESULT 1
US-09-646-651B-2
; Sequence 2, Application US/09646651B
; GENERAL INFORMATION:
;  APPLICANT: Kiesewetter, Stefan
;  APPLICANT:  Kuhn, Eckehard
;  APPLICANT:  Koch-Pelster, Brigitte
;  APPLICANT:  Brunner, Herwig
;  TITLE OF INVENTION: METAL-CONTAINING RIBONUCLEOTIDE POLYPEPTIDES
;  FILE REFERENCE: 206579
;  CURRENT APPLICATION NUMBER: US/09/646,651B
;  CURRENT FILING DATE:  2000-09-13
;  PRIOR APPLICATION NUMBER: PCT/EP98/07722
;  PRIOR FILING DATE: 1998-11-30
;  PRIOR APPLICATION NUMBER: DE 198 11 047.2
;  PRIOR FILING DATE: 1998-03-13
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;  LENGTH: 107
;  TYPE: RNA
;  ORGANISM: Artificial
;  FEATURE:
;  OTHER INFORMATION: Synthetic
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (8)..(8)
;  OTHER INFORMATION: N is nucleotide A, C, U, or G
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (9)..(9)
;  OTHER INFORMATION: N is nucleotide A, C, U, or G
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (10)..(10)
;  OTHER INFORMATION: N is nucleotide A, C, U, or G
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (11)..(11)
;  OTHER INFORMATION: N is nucleotide A, C, U, or G
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (12)..(12)
;  OTHER INFORMATION: N is nucleotide A, C, U, or G
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (14)..(14)
;  OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (20)..(20)
;  OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (21)..(21)
;  OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (22)..(22)
;  OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (23)..(23)
;  OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (24)..(24)
;  OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (25)..(25)
;  OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (28)..(28)
;  OTHER INFORMATION: N is nucleotide A, C, U, or G
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (29)..(29)
;  OTHER INFORMATION: N is nucleotide A, C, U, or G
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (30)..(30)
;  OTHER INFORMATION: N is nucleotide A, C, U, or G
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (34)..(34)
;  OTHER INFORMATION: N is nucleotide A, C, U, or G
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (35)..(35)
;  OTHER INFORMATION: N is nucleotide A, C, U, or G
```

FEATURE:
NAME/KEY: misc_feature
LOCATION: (36)..(36)
OTHER INFORMATION: N is nucleotide A, C, U, or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (37)..(37)
OTHER INFORMATION: N is nucleotide A, C, U, or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (38)..(38)
OTHER INFORMATION: N is nucleotide A, C, U, or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (39)..(39)
OTHER INFORMATION: N is nucleotide A, C, U, or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45)..(45)
OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
FEATURE:
NAME/KEY: misc_feature
LOCATION: (48)..(48)
OTHER INFORMATION: N is nucleotide A, C, U, or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55)..(55)
OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
FEATURE:
NAME/KEY: misc_feature
LOCATION: (56)..(56)
OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
FEATURE:
NAME/KEY: misc_feature
LOCATION: (57)..(57)
OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
FEATURE:
NAME/KEY: misc_feature
LOCATION: (58)..(58)
OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
FEATURE:
NAME/KEY: misc_feature
LOCATION: (59)..(59)
OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
FEATURE:
NAME/KEY: misc_feature
LOCATION: (63)..(63)
OTHER INFORMATION: N is nucleotide A, C, U, or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (66)..(66)
OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
FEATURE:
NAME/KEY: misc_feature
LOCATION: (67)..(67)
OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
FEATURE:
NAME/KEY: misc_feature
LOCATION: (68)..(68)
OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
FEATURE:
NAME/KEY: misc_feature
LOCATION: (69)..(69)
OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
FEATURE:
NAME/KEY: misc_feature
LOCATION: (70)..(70)
OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
FEATURE:
NAME/KEY: misc_feature
LOCATION: (71)..(71)
OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (72)..(72)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (73)..(73)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (74)..(74)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (75)..(75)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (76)..(76)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (77)..(77)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (78)..(78)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (85)..(85)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (86)..(86)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature

Query Match 45.8%; Score 49; DB 27; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.0005;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GGAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNNAAAAANUANAAACAUNNNNNC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNNAAAAANUANAAACAUNNNNNC 60

Qy         61 UUNAGNNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNNUUAGCAG 107
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 UUNAGNNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNNUUAGCAG 107
```

RESULT 2
US-09-646-651C-2
; Sequence 2, Application US/09646651C
; GENERAL INFORMATION:
; APPLICANT: Kiesewetter, Stefan
; APPLICANT: Kuhn, Eckehard
; APPLICANT: Koch-Pelster, Brigitte
; APPLICANT: Brunner, Herwig
; TITLE OF INVENTION: METAL-CONTAINING RIBONUCLEOTIDE POLYPEPTIDES
; FILE REFERENCE: 206579
; CURRENT APPLICATION NUMBER: US/09/646,651C
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: PCT/EP98/07722
; PRIOR FILING DATE: 1998-11-30
; PRIOR APPLICATION NUMBER: DE 198 11 047.2
; PRIOR FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 107
; TYPE: RNA
; ORGANISM: Artificial

; FEATURE:
; OTHER INFORMATION: Synthetic
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8)..(8)
; OTHER INFORMATION: N is nucleotide A, C, U, or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9)..(9)
; OTHER INFORMATION: N is nucleotide A, C, U, or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: N is nucleotide A, C, U, or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: N is nucleotide A, C, U, or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (12)..(12)
; OTHER INFORMATION: N is nucleotide A, C, U, or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14)..(14)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (20)..(20)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (21)..(21)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (22)..(22)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (24)..(24)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (25)..(25)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28)..(28)
; OTHER INFORMATION: N is nucleotide A, C, U, or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29)..(29)
; OTHER INFORMATION: N is nucleotide A, C, U, or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (30)..(30)
; OTHER INFORMATION: N is nucleotide A, C, U, or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (34)..(34)
; OTHER INFORMATION: N is nucleotide A, C, U, or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (35)..(35)
; OTHER INFORMATION: N is nucleotide A, C, U, or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36)..(36)
; OTHER INFORMATION: N is nucleotide A, C, U, or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (37)..(37)
; OTHER INFORMATION: N is nucleotide A, C, U, or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (38)..(38)
; OTHER INFORMATION: N is nucleotide A, C, U, or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (39)..(39)
; OTHER INFORMATION: N is nucleotide A, C, U, or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45)..(45)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (48)..(48)
; OTHER INFORMATION: N is nucleotide A, C, U, or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (55)..(55)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (56)..(56)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (57)..(57)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (58)..(58)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (59)..(59)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (63)..(63)
; OTHER INFORMATION: N is nucleotide A, C, U, or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (66)..(66)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (67)..(67)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (68)..(68)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (69)..(69)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (70)..(70)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (71)..(71)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (72)..(72)
; OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present

```
  FEATURE:
  NAME/KEY: misc_feature
  LOCATION: (73)..(73)
  OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
  FEATURE:
  NAME/KEY: misc_feature
  LOCATION: (74)..(74)
  OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
  FEATURE:
  NAME/KEY: misc_feature
  LOCATION: (75)..(75)
  OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
  FEATURE:
  NAME/KEY: misc_feature
  LOCATION: (76)..(76)
  OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
  FEATURE:
  NAME/KEY: misc_feature
  LOCATION: (77)..(77)
  OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
  FEATURE:
  NAME/KEY: misc_feature
  LOCATION: (78)..(78)
  OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
  FEATURE:
  NAME/KEY: misc_feature
  LOCATION: (85)..(85)
  OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
  FEATURE:
  NAME/KEY: misc_feature
  LOCATION: (86)..(86)
  OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
  FEATURE:
  NAME/KEY: misc_feature

 Query Match             45.8%;  Score 49;  DB 27;  Length 107;
 Best Local Similarity   100.0%;  Pred. No. 0.0005;
 Matches  107;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

y          1 GGAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUUNNNNNNAAAAANUANAAACAUNNNNNC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          1 GGAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUUNNNNNNAAAAANUANAAACAUNNNNNC 60

y         61 UUNAGNNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNNUUAGCAG 107
            |||||||||||||||||||||||||||||||||||||||||||||||||||
b         61 UUNAGNNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNNUUAGCAG 107


ESULT 3
S-09-646-651A-4
 Sequence 4, Application US/09646651A
 GENERAL INFORMATION:
  APPLICANT: Kiesewetter, Stefan
  APPLICANT:  Kuhn, Eckehard
  APPLICANT:  Koch-Pelster, Brigitte
  APPLICANT:  Brunner, Herwig
  TITLE OF INVENTION: METAL-CONTAINING RIBONUCLEOTIDE POLYPEPTIDES
  FILE REFERENCE: 206579
  CURRENT APPLICATION NUMBER: US/09/646,651A
  CURRENT FILING DATE:  2000-09-13
  PRIOR APPLICATION NUMBER: PCT/EP98/07722
  PRIOR FILING DATE: 1998-11-30
  PRIOR APPLICATION NUMBER: DE 198 11 047.2
  PRIOR FILING DATE: 1998-03-13
  NUMBER OF SEQ ID NOS: 4
  SOFTWARE: PatentIn version 3.0
 SEQ ID NO 4
   LENGTH: 298
   TYPE: RNA
   ORGANISM: Unknown
   FEATURE:
   NAME/KEY: misc_feature
   LOCATION: ()..()
;   OTHER INFORMATION: snRNA
US-09-646-651A-4

  Query Match             29.3%;  Score 31.4;  DB 27;  Length 298;
  Best Local Similarity   42.1%;  Pred. No. 58;
  Matches   32;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy         32 UUNNNNNNAAAAAANUANAAACAUNNNNNCUUNAGNNNNNNNNNNNNNNAGAAAUNNNNNNN 91
              ||       |||| || ||||||     ||| ||              ||||||
Db         89 UUUGUACAUAAAAAUAAAAACAUGACUUCUUUAGACACUCCUUCAUUAGAAAUAAAAUAA 148

Qy         92 NNNNNNNNNUUAGCAG 107
                      |||||||
Db        149 AAUAAACUAUUAGCAG 164


RESULT 4
US-09-646-651B-4
; Sequence 4, Application US/09646651B
; GENERAL INFORMATION:
;  APPLICANT: Kiesewetter, Stefan
;  APPLICANT:  Kuhn, Eckehard
;  APPLICANT:  Koch-Pelster, Brigitte
;  APPLICANT:  Brunner, Herwig
;  TITLE OF INVENTION: METAL-CONTAINING RIBONUCLEOTIDE POLYPEPTIDES
;  FILE REFERENCE: 206579
;  CURRENT APPLICATION NUMBER: US/09/646,651B
;  CURRENT FILING DATE:  2000-09-13
;  PRIOR APPLICATION NUMBER: PCT/EP98/07722
;  PRIOR FILING DATE: 1998-11-30
;  PRIOR APPLICATION NUMBER: DE 198 11 047.2
;  PRIOR FILING DATE: 1998-03-13
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
;   LENGTH: 298
;   TYPE: RNA
;   ORGANISM: Unknown
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: ()..()
;   OTHER INFORMATION: snRNA
US-09-646-651B-4

  Query Match             29.3%;  Score 31.4;  DB 27;  Length 298;
  Best Local Similarity   42.1%;  Pred. No. 58;
  Matches   32;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy         32 UUNNNNNNAAAAAANUANAAACAUNNNNNCUUNAGNNNNNNNNNNNNNNAGAAAUNNNNNNN 91
              ||       |||| || ||||||     ||| ||              ||||||
Db         89 UUUGUACAUAAAAAUAAAAACAUGACUUCUUUAGACACUCCUUCAUUAGAAAUAAAAUAA 148

Qy         92 NNNNNNNNNUUAGCAG 107
                      |||||||
Db        149 AAUAAACUAUUAGCAG 164


RESULT 5
US-09-646-651C-4
; Sequence 4, Application US/09646651C
; GENERAL INFORMATION:
;  APPLICANT: Kiesewetter, Stefan
;  APPLICANT:  Kuhn, Eckehard
;  APPLICANT:  Koch-Pelster, Brigitte
;  APPLICANT:  Brunner, Herwig
;  TITLE OF INVENTION: METAL-CONTAINING RIBONUCLEOTIDE POLYPEPTIDES
;  FILE REFERENCE: 206579
;  CURRENT APPLICATION NUMBER: US/09/646,651C
;  CURRENT FILING DATE:  2001-01-16
;  PRIOR APPLICATION NUMBER: PCT/EP98/07722
;  PRIOR FILING DATE: 1998-11-30
;  PRIOR APPLICATION NUMBER: DE 198 11 047.2
;  PRIOR FILING DATE: 1998-03-13
```

```
,  NUMBER OF SEQ ID NOS: 4
,  SOFTWARE: PatentIn version 3.0
  SEQ ID NO 4
    LENGTH: 298
,   TYPE: RNA
,   ORGANISM: Unknown
,   FEATURE:
    NAME/KEY: misc_feature
    LOCATION: ()..()
    OTHER INFORMATION: snRNA
US-09-646-651C-4

  Query Match             29.3%;  Score 31.4;  DB 27;  Length 298;
  Best Local Similarity   42.1%;  Pred. No. 58;
  Matches   32;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy         32 UUNNNNNNAAAAANUANAAACAUNNNNNCUUNAGNNNNNNNNNNNNNNAGAAAUNNNNNNN 91
              ||       |||| || ||||||     ||| ||              ||||||
Db         89 UUUGUACAUAAAAAUAAAAACAUGACUUCUUUAGACACUCCUUCAUUAGAAAUAAAAUAA 148

Qy         92 NNNNNNNNNUUAGCAG 107
                      |||||||
Db        149 AAUAAACUAUUAGCAG 164


RESULT 6
JS-60-466-412-84605
; Sequence 84605, Application US/60466412
, GENERAL INFORMATION:
,  APPLICANT: CARGILL, Michele
,  APPLICANT:  IAKOUBOVA, Olga
,  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
,  TITLE OF INVENTION:  MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001466
;  CURRENT APPLICATION NUMBER: US/60/466,412
;  CURRENT FILING DATE:  2003-04-30
,  NUMBER OF SEQ ID NOS: 429241
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 84605
;   LENGTH: 226307
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(226307)
;   OTHER INFORMATION: n = A,T,C or G
US-60-466-412-84605

  Query Match             27.9%;  Score 29.8;  DB 103;  Length 226307;
  Best Local Similarity   23.4%;  Pred. No. 9.5e+02;
  Matches   25;  Conservative   12;  Mismatches   70;  Indels    0;  Gaps    0;

Qy          1 GGAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNC 60
              |||||       : |: :|       :    ::       ||||  :  || ||:     |
Db      90594 GGAAACATCTATTAATTTGTGACTTTTTACATTGCCCAGAAAATTTGTAAGCATGGTCTC 90653

Qy         61 UUNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAGCAG 107
              :  |               |||||:                 :: | ||
Db      90654 TGAAATAGGCTAAAGAGAAGAAATGGAAAGGAAATGGCCTTTTGTAG 90700


RESULT 7
PCT-US97-14900A-1/c
; Sequence 1, Application PC/TUS9714900A
;  GENERAL INFORMATION:
;    APPLICANT:  THE INSTITUTE FOR GENOMIC RESEARCH
;    APPLICANT:  9712 MEDICAL CENTER DRIVE
;    APPLICANT:  ROCKVILLE, MD 20850
;    APPLICANT:  UNITED STATES OF AMERICA
;    APPLICANT:  MICROBIOLOGY DEPARTMENT
;    APPLICANT:  CHAMPAIGN-URBANA, IL 61801
;    APPLICANT:  UNITED STATES OF AMERICA
;    APPLICANT:  SCHOOL OF MEDICINE
;    APPLICANT:  DEPARTMENT OF MOLECULAR BIOLOGY AND GENETICS
;    APPLICANT:  BALTIMORE, MD 21205
;    APPLICANT:  UNITED STATES OF AMERICA
;    APPLICANT:  APPLICANTS/INVENTORS: Bult, Carol J.
;    APPLICANT:  White, Owen R.
;    APPLICANT:  Smith, Hamilton O.
;    APPLICANT:  Woese, Carl R.
;    APPLICANT:  Venter, J. Craig
;    TITLE OF INVENTION:  Complete Genome Sequence of the Methanogenic
;    TITLE OF INVENTION:  Archaeon, Methanococcus jannaschii
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;      STREET:  1100 New York Avenue, NW, Suite 600
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US97/14900A
;      FILING DATE:  22-AUG-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/024,428
;      FILING DATE:  22-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Eric K. Steffe
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1488.029PC01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1664976 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
PCT-US97-14900A-1

  Query Match             27.5%;  Score 29.4;  DB 1;  Length 1664976;
  Best Local Similarity   24.8%;  Pred. No. 2e+03;
  Matches   26;  Conservative   10;  Mismatches   69;  Indels    0;  Gaps    0;

Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              ||||:     : | | |       |:   :::      ||||| :| |||||:      :
Db     573299 AAAATTAGTATTAGAAGCTATAACTTATTTTAATGCCAAAAAATATAAACATTTCTGGTG 573240

Qy         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUAGCAG 107
                |                ||  |:                  | ||
Db     573239 AAATAATGACAGGCAAAGTTATCTTAGTAGGAGCAGGACCGGGAG 573195


RESULT 8
US-09-692-570-1/c
; Sequence 1, Application US/09692570
; GENERAL INFORMATION:
;  APPLICANT: Bult et al.
;  TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococ
;  TITLE OF INVENTION:  jannaschii
;  FILE REFERENCE: PB275C1
;  CURRENT APPLICATION NUMBER: US/09/692,570
;  CURRENT FILING DATE:  2003-01-14
;  PRIOR APPLICATION NUMBER: US 60/024,428
;  PRIOR FILING DATE: 1996-08-22
;  PRIOR APPLICATION NUMBER: US 08/916,421
;  PRIOR FILING DATE: 1997-08-22
;  NUMBER OF SEQ ID NOS: 20
```

SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1664976
TYPE: DNA
ORGANISM: Methanococcus jannaschii
FEATURE:
NAME/KEY: misc_feature
LOCATION: (28222)..(28222)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (28257)..(28258)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (84773)..(84773)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (103998)..(103998)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (148948)..(148948)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:

```
;   NAME/KEY: misc_feature
;   LOCATION: (779676)..(779676)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (855539)..(855539)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (871619)..(871619)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1084830)..(1084830)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1096846)..(1096846)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1119881)..(1119881)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1130881)..(1130881)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1310988)..(1310988)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1313224)..(1313224)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1349473)..(1349473)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc_feature

  Query Match             27.5%;  Score 29.4;  DB 29;  Length 1664976;
  Best Local Similarity   24.8%;  Pred. No. 2e+03;
  Matches   26;  Conservative   10;  Mismatches   69;  Indels    0;  Gaps    0;

Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              ||||:     : | | |    |:   :::     |||||  :| |||||:     :
Db     573299 AAAATTAGTATTAGAAGCTATAACTTATTTTAATGCCAAAAAATATAAACATTTCTGGTG 573240

Qy         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAGCAG 107
              |              ||  |:                  | ||
Db     573239 AAATAATGACAGGCAAAGTTATCTTAGTAGGAGCAGGACCGGGAG 573195


RESULT 9
US-60-138-831-96/c
; Sequence 96, Application US/60138831
; GENERAL INFORMATION:
;  APPLICANT: Kerlavage, Anthony
;  TITLE OF INVENTION: ISOLATED KINASE PROTEINS, NUCLEIC ACID
;  TITLE OF INVENTION: MOLECULES ENCODING KINASE PROTEINS, AND USES THEREOF
;  FILE REFERENCE: CL000028
;  CURRENT APPLICATION NUMBER: US/60/138,831
;  CURRENT FILING DATE: 1999-06-14
;  NUMBER OF SEQ ID NOS: 620
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 96
;   LENGTH: 609
;   TYPE: DNA
;   ORGANISM: Drosophila
US-60-138-831-96

  Query Match             27.3%;  Score 29.2;  DB 69;  Length 609;
  Best Local Similarity   25.0%;  Pred. No. 2.9e+02;
  Matches   21;  Conservative   13;  Mismatches   50;  Indels    0;  Gaps    0;

Qy          1 GGAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNC 60
              ||||||:     :  :|:       |:   :::      |||||  :| |||  :
Db        263 GGAAAATAAACTTAGTATTTACAATCTTAGTTTTGTACCAAAAAATAAAAATCTGTATAA 204

Qy         61 UUNAGNNNNNNNNNNNNNAGAAAU 84
              :: |                |||:
Db        203 TTAATTTTAAATTCAAATTTAAAT 180


RESULT 10
US-60-140-806-338
; Sequence 338, Application US/60140806
; GENERAL INFORMATION:
;  APPLICANT: Kerlavage, Anthony
;  TITLE OF INVENTION: ISOLATED KINASE PROTEINS, NUCLEIC ACID
;  TITLE OF INVENTION: MOLECULES ENCODING KINASE PROTEINS, AND USES THEREOF
;  FILE REFERENCE: CL000038
;  CURRENT APPLICATION NUMBER: US/60/140,806
;  CURRENT FILING DATE: 1999-06-25
;  NUMBER OF SEQ ID NOS: 1196
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 338
;   LENGTH: 903
;   TYPE: DNA
;   ORGANISM: Drosophila
US-60-140-806-338

  Query Match             27.3%;  Score 29.2;  DB 70;  Length 903;
  Best Local Similarity   25.0%;  Pred. No. 3.2e+02;
  Matches   21;  Conservative   13;  Mismatches   50;  Indels    0;  Gaps    0;

Qy          1 GGAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNC 60
              ||||||:     :  :|:       |:   :::      |||||  :| |||  :
Db        641 GGAAAATAAACTTAGTATTTACAATCTTAGTTTTGTACCAAAAAATAAAAATCTGTATAA 700

Qy         61 UUNAGNNNNNNNNNNNNNAGAAAU 84
              :: |                |||:
Db        701 TTAATTTTAAATTCAAATTTAAAT 724


RESULT 11
US-60-141-856-255
; Sequence 255, Application US/60141856
; GENERAL INFORMATION:
;  APPLICANT: Kerlavage, Anthony
;  TITLE OF INVENTION: ISOLATED KINASE PROTEINS, NUCLEIC ACID
;  TITLE OF INVENTION: MOLECULES ENCODING KINASE PROTEINS, AND USES THEREOF
;  FILE REFERENCE: CL000044
;  CURRENT APPLICATION NUMBER: US/60/141,856
;  CURRENT FILING DATE: 1999-07-01
;  NUMBER OF SEQ ID NOS: 617
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 255
;   LENGTH: 2334
;   TYPE: DNA
;   ORGANISM: Drosophila
US-60-141-856-255

  Query Match             27.3%;  Score 29.2;  DB 70;  Length 2334;
  Best Local Similarity   25.0%;  Pred. No. 4.2e+02;
  Matches   21;  Conservative   13;  Mismatches   50;  Indels    0;  Gaps    0;

Qy          1 GGAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNC 60
              ||||||:     :  :|:       |:   :::      |||||  :| |||  :
Db       2072 GGAAAATAAACTTAGTATTTACAATCTTAGTTTTGTACCAAAAAATAAAAATCTGTATAA 2131

Qy         61 UUNAGNNNNNNNNNNNNNAGAAAU 84
              :: |                |||:
Db       2132 TTAATTTTAAATTCAAATTTAAAT 2155
```

ESULT 12
S-60-173-464-20990
Sequence 20990, Application US/60173464
GENERAL INFORMATION:
APPLICANT: Li, Peter W.D.
TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CL000173
CURRENT APPLICATION NUMBER: US/60/173,464
CURRENT FILING DATE: 1999-12-29
NUMBER OF SEQ ID NOS: 30269
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 20990
LENGTH: 3020
TYPE: DNA
ORGANISM: Drosophila
S-60-173-464-20990

Query Match 27.3%; Score 29.2; DB 73; Length 3020;
Best Local Similarity 25.0%; Pred. No. 4.5e+02;
Matches 21; Conservative 13; Mismatches 50; Indels 0; Gaps 0;

```
Qy          1 GGAAAAUNNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNC 60
              ||||||:     :  :|:       |:   :::      ||||| :| |||  :
Db        769 GGAAAATAAACTTAGTATTTACAATCTTAGTTTTGTACCAAAAAATAAAAATCTGTATAA 828

Qy         61 UUNAGNNNNNNNNNNNNNAGAAAU 84
              :: |                |||:
Db        829 TTAATTTTAAATTCAAATTTAAAT 852
```

ESULT 13
S-60-191-637-25467
Sequence 25467, Application US/60191637
GENERAL INFORMATION:
APPLICANT: Venter, J. Craig
TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING
TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: CL000392
CURRENT APPLICATION NUMBER: US/60/191,637
CURRENT FILING DATE: 2000-03-23
NUMBER OF SEQ ID NOS: 42660
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 25467
LENGTH: 3020
TYPE: DNA
ORGANISM: DROSOPHILA
S-60-191-637-25467

Query Match 27.3%; Score 29.2; DB 75; Length 3020;
Best Local Similarity 25.0%; Pred. No. 4.5e+02;
Matches 21; Conservative 13; Mismatches 50; Indels 0; Gaps 0;

```
Qy          1 GGAAAAUNNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNC 60
              ||||||:     :  :|:       |:   :::      ||||| :| |||  :
Db        769 GGAAAATAAACTTAGTATTTACAATCTTAGTTTTGTACCAAAAAATAAAAATCTGTATAA 828

Qy         61 UUNAGNNNNNNNNNNNNNAGAAAU 84
              :: |                |||:
Db        829 TTAATTTTAAATTCAAATTTAAAT 852
```

ESULT 14
S-60-191-681-20076
Sequence 20076, Application US/60191681
GENERAL INFORMATION:
APPLICANT: Li, Peter, W.D.
TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING ISOLATED DROSOPHILA PROTEINS AND
; TITLE OF INVENTION: USES THEREOF.
; FILE REFERENCE: cl000390
; CURRENT APPLICATION NUMBER: US/60/191,681
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 30973
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20076
; LENGTH: 3020
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-60-191-681-20076

Query Match 27.3%; Score 29.2; DB 75; Length 3020;
Best Local Similarity 25.0%; Pred. No. 4.5e+02;
Matches 21; Conservative 13; Mismatches 50; Indels 0; Gaps 0;

```
Qy          1 GGAAAAUNNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNC 60
              ||||||:     :  :|:       |:   :::      ||||| :| |||  :
Db        769 GGAAAATAAACTTAGTATTTACAATCTTAGTTTTGTACCAAAAAATAAAAATCTGTATAA 828

Qy         61 UUNAGNNNNNNNNNNNNNAGAAAU 84
              :: |                |||:
Db        829 TTAATTTTAAATTCAAATTTAAAT 852
```

RESULT 15
US-60-142-844-270
; Sequence 270, Application US/60142844
; GENERAL INFORMATION:
; APPLICANT: Kerlavage, Anthony
; TITLE OF INVENTION: ISOLATED KINASE PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES ENCODING KINASE PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000049
; CURRENT APPLICATION NUMBER: US/60/142,844
; CURRENT FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 730
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 270
; LENGTH: 3238
; TYPE: DNA
; ORGANISM: Drosophila
US-60-142-844-270

Query Match 27.3%; Score 29.2; DB 70; Length 3238;
Best Local Similarity 25.0%; Pred. No. 4.5e+02;
Matches 21; Conservative 13; Mismatches 50; Indels 0; Gaps 0;

```
Qy          1 GGAAAAUNNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNC 60
              ||||||:     :  :|:       |:   :::      ||||| :| |||  :
Db       2073 GGAAAATAAACTTAGTATTTACAATCTTAGTTTTGTACCAAAAAATAAAAATCTGTATAA 2132

Qy         61 UUNAGNNNNNNNNNNNNNAGAAAU 84
              :: |                |||:
Db       2133 TTAATTTTAAATTCAAATTTAAAT 2156
```

Search completed: March 8, 2004, 07:15:09
Job time : 3141 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 8, 2004, 05:20:29 ; Search time 44 Seconds
(without alignments)
597.965 Million cell updates/sec

Title: US-09-646-651C-2
Perfect score: 107
Sequence: 1 ggaaaaunnnnnunauaugn.....nnnnnnnnnnnnnnuuagcag 107

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 517230 seqs, 122946014 residues

Total number of hits satisfying chosen parameters: 1034460

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	27.4	25.6	524	6	US-10-779-543-22280	Sequence 22280, A
	2	27.4	25.6	761	6	US-10-779-543-22493	Sequence 22493, A
	3	27.4	25.6	769	6	US-10-779-543-22418	Sequence 22418, A
	4	27.4	25.6	789	6	US-10-779-543-22415	Sequence 22415, A
	5	27.4	25.6	929	6	US-10-779-543-23307	Sequence 23307, A
	6	24.6	23.0	3552	6	US-10-775-972-126	Sequence 126, App
C	7	24.6	23.0	1790242	6	US-10-767-471-10805	Sequence 10805, A
C	8	24.4	22.8	500	1	PCT-US04-02000-361	Sequence 361, App
	9	24.2	22.6	1000	1	PCT-US04-02000-172	Sequence 172, App
C	10	23.8	22.2	184809	6	US-10-767-471-10627	Sequence 10627, A
C	11	23.6	22.1	104000	6	US-10-673-523-14	Sequence 14, Appl
C	12	23.6	22.1	138434	6	US-10-767-471-10732	Sequence 10732, A
C	13	23.4	21.9	1908	6	US-10-468-591A-8	Sequence 8, Appli
	14	23.4	21.9	3258	5	US-09-830-230A-324	Sequence 324, App
	15	23.4	21.9	3354	5	US-09-830-230A-323	Sequence 323, App
	16	23.4	21.9	12751	6	US-10-767-471-10612	Sequence 10612, A
C	17	23.4	21.9	138244	6	US-10-767-471-10702	Sequence 10702, A
C	18	23.2	21.7	1119	6	US-10-091-007A-197	Sequence 197, App
C	19	23.2	21.7	116733	6	US-10-767-471-10741	Sequence 10741, A
C	20	23.2	21.7	130030	6	US-10-767-471-10854	Sequence 10854, A
C	21	23.2	21.7	165097	6	US-10-765-790-77	Sequence 77, Appl
C	22	23.2	21.7	166020	6	US-10-767-471-10806	Sequence 10806, A
	23	23.2	21.7	193757	6	US-10-767-471-10614	Sequence 10614, A
C	24	23.2	21.7	546025	6	US-10-767-471-10717	Sequence 10717, A
	25	23.2	21.7	1790242	6	US-10-767-471-10805	Sequence 10805, A
	26	23	21.5	549	6	US-10-767-701-30635	Sequence 30635, A
	27	23	21.5	1092	6	US-10-417-884A-3103	Sequence 3103, Ap
	28	23	21.5	388939	6	US-10-417-375A-4	Sequence 4, Appli
C	29	23	21.5	495635	6	US-10-765-790-12	Sequence 12, Appl
C	30	23	21.5	705636	6	US-10-765-790-30	Sequence 30, Appl
	31	22.8	21.3	77483	6	US-10-767-471-10844	Sequence 10844, A
C	32	22.6	21.1	580	6	US-10-779-543-20980	Sequence 20980, A
	33	22.6	21.1	67253	6	US-10-765-790-88	Sequence 88, Appl
	34	22.6	21.1	102090	6	US-10-767-471-10823	Sequence 10823, A
C	35	22.6	21.1	102090	6	US-10-767-471-10823	Sequence 10823, A
	36	22.4	20.9	411	6	US-10-417-884A-1919	Sequence 1919, Ap
	37	22.4	20.9	526	6	US-10-779-543-20015	Sequence 20015, A
	38	22.4	20.9	534	6	US-10-779-543-19540	Sequence 19540, A
	39	22.4	20.9	631	6	US-10-779-543-22233	Sequence 22233, A
C	40	22.4	20.9	822	6	US-10-779-543-22331	Sequence 22331, A
	41	22.4	20.9	873	6	US-10-417-884A-2829	Sequence 2829, Ap
C	42	22.4	20.9	38142	6	US-10-417-375A-55	Sequence 55, Appl
C	43	22.4	20.9	101001	6	US-10-767-471-10650	Sequence 10650, A
C	44	22.4	20.9	170546	6	US-10-767-471-10703	Sequence 10703, A
C	45	22.4	20.9	803755	6	US-10-767-471-10700	Sequence 10700, A

ALIGNMENTS

```
RESULT 1
US-10-779-543-22280
; Sequence 22280, Application US/10779543
; GENERAL INFORMATION:
;  APPLICANT: Williams et al
;  TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
;  TITLE OF INVENTION:  IN CANCEROUS CELLS AND THEIR METHODS OF USE II
;  FILE REFERENCE: 2300-21302
;  CURRENT APPLICATION NUMBER: US/10/779,543
;  CURRENT FILING DATE:  2004-02-12
;  PRIOR APPLICATION NUMBER: 10/076,555
;  PRIOR FILING DATE: 2002-02-15
;  PRIOR APPLICATION NUMBER: 09/217,471
;  PRIOR FILING DATE: 1998-12-21
;  PRIOR APPLICATION NUMBER: 60/068,755
;  PRIOR FILING DATE: 1997-12-23
;  PRIOR APPLICATION NUMBER: 60/080,664
;  PRIOR FILING DATE: 1998-04-03
;  PRIOR APPLICATION NUMBER: 60/105,234
;  PRIOR FILING DATE: 1998-10-21
;  PRIOR APPLICATION NUMBER: 09/297,648
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: PCT/US99/01619
;  PRIOR FILING DATE: 1999-01-28
;  PRIOR APPLICATION NUMBER: 60/072,910
;  PRIOR FILING DATE: 1998-01-28
;  PRIOR APPLICATION NUMBER: 60/075,954
;  PRIOR FILING DATE: 1998-02-24
;  PRIOR APPLICATION NUMBER: 60/080,114
;  PRIOR FILING DATE: 1998-03-31
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 23767
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 22280
;   LENGTH: 524
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 479, 484, 486, 487, 496, 501, 505, 506, 509, 511, 513, 514,
;   LOCATION: 515, 516, 521, 523
;   OTHER INFORMATION: n = A,T,C or G
US-10-779-543-22280

  Query Match           25.6%;  Score 27.4;  DB 6;  Length 524;
  Best Local Similarity  24.3%;  Pred. No. 5.1;
  Matches   25;  Conservative    9;  Mismatches   69;  Indels    0;  Gaps    0;

Qy          2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAAANUANAAACAUNNNNNCU 61
```

```
              |||||      : | |         :    :       ||||  :| |||| :     |:
Db        232 GAAAACTCTTGTTAGACATGATGTTTCCTATACAGAAGAAAAGATACAAACTTGATTGCT 291

Qy         62 UNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAG 104
              : ||              ||| |:                 : ||
Db        292 TCAGTAAAAACACTTTTAGAGATTGTGTTTCCTTTTTTGTGAG 334
```

RESULT 2
US-10-779-543-22493
; Sequence 22493, Application US/10779543
; GENERAL INFORMATION:
; APPLICANT: Williams et al
; TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
; TITLE OF INVENTION: IN CANCEROUS CELLS AND THEIR METHODS OF USE II
; FILE REFERENCE: 2300-21302
; CURRENT APPLICATION NUMBER: US/10/779,543
; CURRENT FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 10/076,555
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 09/217,471
; PRIOR FILING DATE: 1998-12-21
; PRIOR APPLICATION NUMBER: 60/068,755
; PRIOR FILING DATE: 1997-12-23
; PRIOR APPLICATION NUMBER: 60/080,664
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/105,234
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: 09/297,648
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: PCT/US99/01619
; PRIOR FILING DATE: 1999-01-28
; PRIOR APPLICATION NUMBER: 60/072,910
; PRIOR FILING DATE: 1998-01-28
; PRIOR APPLICATION NUMBER: 60/075,954
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/080,114
; PRIOR FILING DATE: 1998-03-31
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 23767
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 22493
; LENGTH: 761
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 534, 596, 615, 633, 641, 648, 656, 710, 729, 736, 744, 748,
; LOCATION: 749, 751, 752
; OTHER INFORMATION: n = A,T,C or G
US-10-779-543-22493

Query Match 25.6%; Score 27.4; DB 6; Length 761;
Best Local Similarity 24.3%; Pred. No. 5.4;
Matches 25; Conservative 9; Mismatches 69; Indels 0; Gaps 0;

```
Qy          2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNNAAAAANUANAAACAUNNNNNCU 61
              |||||      : | |         :    :       ||||  :| |||| :     |:
Db        237 GAAAACTCTTGTTAGACATGATGTTTCCTATACAGAAGAAAAGATACAAACTTGATTGCT 296

Qy         62 UNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAG 104
              : ||              ||| |:                 : ||
Db        297 TCAGTAAAAACACTTTTAGAGATTGTGTTTCCTTTTTTGTGAG 339
```

RESULT 3
US-10-779-543-22418
; Sequence 22418, Application US/10779543
; GENERAL INFORMATION:
; APPLICANT: Williams et al
; TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
; TITLE OF INVENTION: IN CANCEROUS CELLS AND THEIR METHODS OF USE II
; FILE REFERENCE: 2300-21302
; CURRENT APPLICATION NUMBER: US/10/779,543
; CURRENT FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 10/076,555
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 09/217,471
; PRIOR FILING DATE: 1998-12-21
; PRIOR APPLICATION NUMBER: 60/068,755
; PRIOR FILING DATE: 1997-12-23
; PRIOR APPLICATION NUMBER: 60/080,664
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/105,234
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: 09/297,648
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: PCT/US99/01619
; PRIOR FILING DATE: 1999-01-28
; PRIOR APPLICATION NUMBER: 60/072,910
; PRIOR FILING DATE: 1998-01-28
; PRIOR APPLICATION NUMBER: 60/075,954
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/080,114
; PRIOR FILING DATE: 1998-03-31
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 23767
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 22418
; LENGTH: 769
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2, 3, 4, 8, 19, 21, 37, 40, 44, 45, 54, 57, 65, 66, 71, 500,
; LOCATION: 566, 574, 575, 576, 597, 604, 617, 633, 640, 656, 660, 680,
; LOCATION: 684, 694, 709, 716, 720, 730, 734, 750, 755, 762, 763, 766
; OTHER INFORMATION: n = A,T,C or G
US-10-779-543-22418

Query Match 25.6%; Score 27.4; DB 6; Length 769;
Best Local Similarity 24.3%; Pred. No. 5.4;
Matches 25; Conservative 9; Mismatches 69; Indels 0; Gaps 0;

```
Qy          2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNNAAAAANUANAAACAUNNNNNCU 61
              |||||      : | |         :    :       ||||  :| |||| :     |:
Db         88 GAAAACTCTTGTTAGACATGATGTTTCCTATACAGAAGAAAAGATACAAACTTGATTGCT 147

Qy         62 UNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAG 104
              : ||              ||| |:                 : ||
Db        148 TCAGTAAAAACACTTTTAGAGATTGTGTTTCCTTTTTTGTGAG 190
```

RESULT 4
US-10-779-543-22415
; Sequence 22415, Application US/10779543
; GENERAL INFORMATION:
; APPLICANT: Williams et al
; TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
; TITLE OF INVENTION: IN CANCEROUS CELLS AND THEIR METHODS OF USE II
; FILE REFERENCE: 2300-21302
; CURRENT APPLICATION NUMBER: US/10/779,543
; CURRENT FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 10/076,555
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 09/217,471
; PRIOR FILING DATE: 1998-12-21
; PRIOR APPLICATION NUMBER: 60/068,755
; PRIOR FILING DATE: 1997-12-23
; PRIOR APPLICATION NUMBER: 60/080,664
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/105,234
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: 09/297,648
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: PCT/US99/01619

```
 PRIOR FILING DATE: 1999-01-28
 PRIOR APPLICATION NUMBER: 60/072,910
 PRIOR FILING DATE: 1998-01-28
 PRIOR APPLICATION NUMBER: 60/075,954
 PRIOR FILING DATE: 1998-02-24
 PRIOR APPLICATION NUMBER: 60/080,114
 PRIOR FILING DATE: 1998-03-31
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 23767
 SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 22415
  LENGTH: 789
  TYPE: DNA
  ORGANISM: Homo sapiens
  FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 5, 7, 9, 12, 15, 17, 19, 20, 27, 33, 46, 62, 63, 64, 66,
  LOCATION: 67, 74, 76, 509, 613, 645, 651, 655, 666, 678, 715, 718,
  LOCATION: 729, 742, 772, 776, 779, 780, 781, 783, 786, 787, 788
  OTHER INFORMATION: n = A,T,C or G
S-10-779-543-22415

 Query Match             25.6%;  Score 27.4;  DB 6;  Length 789;
 Best Local Similarity   24.3%;  Pred. No. 5.4;
 Matches   25;  Conservative    9;  Mismatches   69;  Indels    0;  Gaps    0;

y          2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCU 61
             |||||      : | |         :    :       ||||  :| |||| :     |:
b         97 GAAAACTCTTGTTAGACATGATGTTTCCTATACAGAAGAAAAGATACAAACTTGATTGCT 156

y         62 UNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUAG 104
             : ||              ||| |:                : ||
b        157 TCAGTAAAAACACTTTTAGAGATTGTGTTTCCTTTTTTGTGAG 199


ESULT 5
S-10-779-543-23307
 Sequence 23307, Application US/10779543
 GENERAL INFORMATION:
  APPLICANT: Williams et al
  TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
  TITLE OF INVENTION:  IN CANCEROUS CELLS AND THEIR METHODS OF USE II
  FILE REFERENCE: 2300-21302
  CURRENT APPLICATION NUMBER: US/10/779,543
  CURRENT FILING DATE:  2004-02-12
  PRIOR APPLICATION NUMBER: 10/076,555
  PRIOR FILING DATE: 2002-02-15
  PRIOR APPLICATION NUMBER: 09/217,471
  PRIOR FILING DATE: 1998-12-21
  PRIOR APPLICATION NUMBER: 60/068,755
  PRIOR FILING DATE: 1997-12-23
  PRIOR APPLICATION NUMBER: 60/080,664
  PRIOR FILING DATE: 1998-04-03
  PRIOR APPLICATION NUMBER: 60/105,234
  PRIOR FILING DATE: 1998-10-21
  PRIOR APPLICATION NUMBER: 09/297,648
  PRIOR FILING DATE: 2000-04-10
  PRIOR APPLICATION NUMBER: PCT/US99/01619
  PRIOR FILING DATE: 1999-01-28
  PRIOR APPLICATION NUMBER: 60/072,910
  PRIOR FILING DATE: 1998-01-28
  PRIOR APPLICATION NUMBER: 60/075,954
  PRIOR FILING DATE: 1998-02-24
  PRIOR APPLICATION NUMBER: 60/080,114
  PRIOR FILING DATE: 1998-03-31
  Remaining Prior Application data removed - See File Wrapper or PALM.
  NUMBER OF SEQ ID NOS: 23767
  SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 23307
   LENGTH: 929
   TYPE: DNA
   ORGANISM: Homo sapiens
   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 649, 785, 791, 795, 806, 818, 855, 858, 869, 882, 912, 916,
;   LOCATION: 919, 920, 921, 923, 926, 927, 928
;   OTHER INFORMATION: n = A,T,C or G
US-10-779-543-23307

  Query Match             25.6%;  Score 27.4;  DB 6;  Length 929;
  Best Local Similarity   24.3%;  Pred. No. 5.6;
  Matches   25;  Conservative    9;  Mismatches   69;  Indels    0;  Gaps    0;

Qy         2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCU 61
             |||||      : | |         :    :       ||||  :| |||| :     |:
Db       237 GAAAACTCTTGTTAGACATGATGTTTCCTATACAGAAGAAAAGATACAAACTTGATTGCT 296

Qy        62 UNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUAG 104
             : ||              ||| |:                : ||
Db       297 TCAGTAAAAACACTTTTAGAGATTGTGTTTCCTTTTTTGTGAG 339


RESULT 6
US-10-775-972-126
; Sequence 126, Application US/10775972
; GENERAL INFORMATION:
;  APPLICANT: Henderson, Robert A.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C21
;  CURRENT APPLICATION NUMBER: US/10/775,972
;  CURRENT FILING DATE:  2004-02-10
;  NUMBER OF SEQ ID NOS: 563
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 126
;   LENGTH: 3552
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-775-972-126

  Query Match             23.0%;  Score 24.6;  DB 6;  Length 3552;
  Best Local Similarity   27.2%;  Pred. No. 43;
  Matches   22;  Conservative    8;  Mismatches   51;  Indels    0;  Gaps    0;

Qy         3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
             || |:     :  :|:|      |:     :       |||  | |||||       ::
Db      3081 AATATCACATTATTATGTATTCACTTTAAGTGATAGTTTAAAAAATAAACAAGAAATATT 3140

Qy        63 NAGNNNNNNNNNNNNNAGAAA 83
             ||               |||||
Db      3141 GAGTATCACTATGTGAAGAAA 3161


RESULT 7
US-10-767-471-10805/c
; Sequence 10805, Application US/10767471
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001505
;  CURRENT APPLICATION NUMBER: US/10/767,471
;  CURRENT FILING DATE:  2004-01-30
;  NUMBER OF SEQ ID NOS: 50231
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10805
;   LENGTH: 1790242
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1790242)
;   OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-767-471-10805
```

Query Match 23.0%; Score 24.6; DB 6; Length 1790242;
Best Local Similarity 22.9%; Pred. No. 73;
Matches 24; Conservative 9; Mismatches 72; Indels 0; Gaps 0;

```
Qy          3 AAAAAUNNNNNUNAUAUGNNNNNNCUNNNNUUUNNNNNNNAAAAAANUANAAACAUNNNNNCUU 62
              ||||:       |: :|             ::      |||||  |  ||||:      ::
Db    1427735 AAAATTTCCTCCATTTGTGGTTAAATGTCTTTTACCAAAAAAAGATGAACATTGAGAGTT 1427676

Qy         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUAGCAG 107
              ||             ||                 :  ||||
Db    1427675 CAGTTCTCATTTATGTTTTAAGGGGTCTATAAAAATTATGTGCAG 1427631
```

RESULT 8
PCT-US04-02000-361/c
; Sequence 361, Application PC/TUS0402000
; GENERAL INFORMATION:
; APPLICANT: Science Applications International Corporation
; TITLE OF INVENTION: Method and System for Identifying Biological Entities in
; TITLE OF INVENTION: Biological and Environmental Samples
; FILE REFERENCE: 36609-183264 (SAIC0087-PCT)
; CURRENT APPLICATION NUMBER: PCT/US04/02000
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US 60/441,745
; PRIOR FILING DATE: 2003-01-23
; PRIOR APPLICATION NUMBER: US 60/441,806
; PRIOR FILING DATE: 2003-01-23
; NUMBER OF SEQ ID NOS: 3242
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 361
; LENGTH: 500
; TYPE: DNA
; ORGANISM: Clostridium perfringens
PCT-US04-02000-361

Query Match 22.8%; Score 24.4; DB 1; Length 500;
Best Local Similarity 19.0%; Pred. No. 36;
Matches 19; Conservative 12; Mismatches 69; Indels 0; Gaps 0;

```
Qy          4 AAAAUNNNNNUNAUAUGNNNNNNCUNNNNUUUNNNNNNNAAAAAANUANAAACAUNNNNNCUUN 63
              | |:     : | |:          :::      |||   :| ||| |:       :
Db        476 ATATTTTCATAAAATCCTATTTTATATTTTTCTTTTAAATGTTAAAAATATAATTATATA 417

Qy         64 AGNNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUA 103
              ||                |||||:                ::|
Db        416 AGTTCTATTATTTAATGAAATTTTGTGTTAAATTAGATTA 377
```

RESULT 9
PCT-US04-02000-172
; Sequence 172, Application PC/TUS0402000
; GENERAL INFORMATION:
; APPLICANT: Science Applications International Corporation
; TITLE OF INVENTION: Method and System for Identifying Biological Entities in
; TITLE OF INVENTION: Biological and Environmental Samples
; FILE REFERENCE: 36609-183264 (SAIC0087-PCT)
; CURRENT APPLICATION NUMBER: PCT/US04/02000
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US 60/441,745
; PRIOR FILING DATE: 2003-01-23
; PRIOR APPLICATION NUMBER: US 60/441,806
; PRIOR FILING DATE: 2003-01-23
; NUMBER OF SEQ ID NOS: 3242
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 172
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Clostridium perfringens
PCT-US04-02000-172

Query Match 22.6%; Score 24.2; DB 1; Length 1000;
Best Local Similarity 19.4%; Pred. No. 46;
Matches 20; Conservative 12; Mismatches 71; Indels 0; Gaps 0;

```
Qy          4 AAAUNNNNNUNAUAUGNNNNNNCUNNNNUUUNNNNNNNAAAAAANUANAAACAUNNNNNCUUN 63
              ||        : |:|:|       |:   :::       || || :| ||   :       :
Db        710 AACAGTTGGTTATATGTAGGAGCTCCCTTTAAGTTTAATAAATAAAATATTTTTGTTTAG 769

Qy         64 AGNNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAGCA 106
              ||                  |||                  :: | |
Db        770 AGTATATTTTCTTAATTAAAAGAGCTTAAAAATAAGTTTTGTA 812
```

RESULT 10
US-10-767-471-10627/c
; Sequence 10627, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10627
; LENGTH: 184809
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(184809)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-767-471-10627

Query Match 22.2%; Score 23.8; DB 6; Length 184809;
Best Local Similarity 20.8%; Pred. No. 1.2e+02;
Matches 21; Conservative 10; Mismatches 70; Indels 0; Gaps 0;

```
Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNNAAAAAANUANAAACAUNNNNNCUU 62
              | ||       : |: :|            :::       |||||:| ||| |       |::
Db      75589 ATAAGTACAATTATGTGTCAATTTAAATTTTAATTTTAAAAATTAAAAAAAAGAAGGCTT 75530

Qy         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUA 103
              |                  |||                :|
Db      75529 TGGGAGCCCAACACATCCAAAGAGGAAACCATCTTCAGGTA 75489
```

RESULT 11
US-10-673-523-14/c
; Sequence 14, Application US/10673523
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOLIPID SCRAMBLASE 4 EXPRESSION
; FILE REFERENCE: RTS-0334
; CURRENT APPLICATION NUMBER: US/10/673,523
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: US/10/012,984
; PRIOR FILING DATE: 2001-12-04
; NUMBER OF SEQ ID NOS: 92
; SEQ ID NO 14
; LENGTH: 104000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 14992
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 14993
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 14994

```
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 14995
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 14996
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 14997
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 14998
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 14999
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 15000
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 15001
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 15002
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 15003
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 15004
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 15005
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 15006
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 15007
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 15008
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 15009
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 15010
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 15011
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 15012
OTHER INFORMATION: unknown
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: 15013
;  OTHER INFORMATION: unknown
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: 15014
;  OTHER INFORMATION: unknown
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: 15015
;  OTHER INFORMATION: unknown
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: 15016
;  OTHER INFORMATION: unknown
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: 15017
;  OTHER INFORMATION: unknown
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: 15018
;  OTHER INFORMATION: unknown
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: 15019
;  OTHER INFORMATION: unknown
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: 15020
;  OTHER INFORMATION: unknown
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: 15021
;  OTHER INFORMATION: unknown
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: 15022
;  OTHER INFORMATION: unknown
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: 15023
;  OTHER INFORMATION: unknown
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: 15024
;  OTHER INFORMATION: unknown
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: 15025
;  OTHER INFORMATION: unknown
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: 15026
;  OTHER INFORMATION: unknown
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: 15027
;  OTHER INFORMATION: unknown
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: 15028
;  OTHER INFORMATION: unknown
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: 15029
;  OTHER INFORMATION: unknown
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: 15030
;  OTHER INFORMATION: unknown
;  FEATURE:
```

```
;   NAME/KEY: unsure
;   LOCATION: 15031
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 15032
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 15033
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 15034
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 15035
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 15036
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 15037
;   OTHER INFORMATION: unknown
;   FEATURE:
;   NAME/KEY: unsure

  Query Match             22.1%;  Score 23.6;  DB 6;  Length 104000;
  Best Local Similarity   23.1%;  Pred. No. 1.3e+02;
  Matches   24;  Conservative    8;  Mismatches   72;  Indels    0;  Gaps    0;

Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              ||||      : | |                ::      ||||  :  ||||       |::
Db      13486 AAAAAGAAAATCAAAAAAGCTAAAAATTGTTGCTTTGAAAAGATCAAAACTGACAAACTT 13427

Qy         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUAGCA 106
                |           |||||              ::| ||
Db      13426 TTGGATACACTAATAAAGAAAAAGAGAAAAGATTCAAATTACCA 13383


RESULT 12
US-10-767-471-10732/c
; Sequence 10732, Application US/10767471
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001505
;  CURRENT APPLICATION NUMBER: US/10/767,471
;  CURRENT FILING DATE:  2004-01-30
;  NUMBER OF SEQ ID NOS: 50231
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10732
;   LENGTH: 138434
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-767-471-10732

  Query Match             22.1%;  Score 23.6;  DB 6;  Length 138434;
  Best Local Similarity   23.1%;  Pred. No. 1.3e+02;
  Matches   24;  Conservative    8;  Mismatches   72;  Indels    0;  Gaps    0;

Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              |||       : |:|        |    ::      |||||  | ||| |:     |::
Db     123972 AAACAGTATATTATACCATTATACCATATTAGTGACTAAAAACGAAAAATATATGATCTT 123913

Qy         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUAGCA 106
                ||              |||  |                :  | |
Db     123912 GAGGCACAAAAGCATATGATAAAAGTTTTTTTTGCTATTATGAA 123869


RESULT 13
US-10-468-591A-8/c
; Sequence 8, Application US/10468591A
; GENERAL INFORMATION:
;  APPLICANT: Sahin, Ugur
;  APPLICANT:  Tureci, Ozlem
;  APPLICANT:  Ludewig, Burkhard
;  TITLE OF INVENTION: Method for identifying biologically active structures of microbia
;  TITLE OF INVENTION:  pathogens
;  FILE REFERENCE: 03-757
;  CURRENT APPLICATION NUMBER: US/10/468,591A
;  CURRENT FILING DATE:  2004-02-06
;  PRIOR APPLICATION NUMBER: PCT/EP02/01909
;  PRIOR FILING DATE: 2002-02-22
;  PRIOR APPLICATION NUMBER: DE 10108626.1
;  PRIOR FILING DATE: 2001-02-22
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
;   LENGTH: 1908
;   TYPE: DNA
;   ORGANISM: Vaccinia virus
US-10-468-591A-8

  Query Match             21.9%;  Score 23.4;  DB 6;  Length 1908;
  Best Local Similarity   29.5%;  Pred. No. 85;
  Matches   18;  Conservative    9;  Mismatches   34;  Indels    0;  Gaps    0;

Qy          2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCU 61
              || ||:       |:| |        :   :::      ||||  :| |||||:     |
Db       1396 GATAATAGGGCATATAGGTCATCATTATATTTGGATTAAAAATTTAGAAACATATACGCG 1337

Qy         62 U 62
              :
Db       1336 T 1336


RESULT 14
US-09-830-230A-324
; Sequence 324, Application US/09830230A
; GENERAL INFORMATION:
;  APPLICANT: Human Genome Sciences, Inc.
;  TITLE OF INVENTION: Lyme Disease Vaccines
;  FILE REFERENCE: PB481US
;  CURRENT APPLICATION NUMBER: US/09/830,230A
;  CURRENT FILING DATE:  2001-09-27
;  PRIOR APPLICATION NUMBER: PCT/US98/12718
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/057,483
;  PRIOR FILING DATE: 1997-09-03
;  PRIOR APPLICATION NUMBER: 60/053,344
;  PRIOR FILING DATE: 1997-07-22
;  PRIOR APPLICATION NUMBER: 60/053,377
;  PRIOR FILING DATE: 1997-07-22
;  PRIOR APPLICATION NUMBER: 60/050,359
;  PRIOR FILING DATE: 1997-06-20
;  NUMBER OF SEQ ID NOS: 756
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 324
;   LENGTH: 3258
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-830-230A-324

  Query Match             21.9%;  Score 23.4;  DB 5;  Length 3258;
  Best Local Similarity   21.2%;  Pred. No. 92;
  Matches   21;  Conservative    9;  Mismatches   69;  Indels    0;  Gaps    0;

Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              ||||      : |:|          :   :::        || :| |||||       |
Db       2324 AAAAAGCACTTCATAACAAAGGAATAGCTTTAATGATGCTAAATAAAAACAAAAAAGCAA 2383
```

```
Qy          63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNU 101
               ||               || ||:                :
Db        2384 TAGAATCTTTTGAGAAAGCAATACAAATTGATAAAAATT 2422


ESULT 15
S-09-830-230A-323
 Sequence 323, Application US/09830230A
 GENERAL INFORMATION:
  APPLICANT: Human Genome Sciences, Inc.
  TITLE OF INVENTION: Lyme Disease Vaccines
  FILE REFERENCE: PB481US
  CURRENT APPLICATION NUMBER: US/09/830,230A
  CURRENT FILING DATE:  2001-09-27
  PRIOR APPLICATION NUMBER: PCT/US98/12718
  PRIOR FILING DATE: 1998-06-18
  PRIOR APPLICATION NUMBER: 60/057,483
  PRIOR FILING DATE: 1997-09-03
  PRIOR APPLICATION NUMBER: 60/053,344
  PRIOR FILING DATE: 1997-07-22
  PRIOR APPLICATION NUMBER: 60/053,377
  PRIOR FILING DATE: 1997-07-22
  PRIOR APPLICATION NUMBER: 60/050,359
  PRIOR FILING DATE: 1997-06-20
  NUMBER OF SEQ ID NOS: 756
  SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 323
   LENGTH: 3354
   TYPE: DNA
   ORGANISM: Homo sapiens
S-09-830-230A-323

 Query Match             21.9%;  Score 23.4;  DB 5;  Length 3354;
 Best Local Similarity   21.2%;  Pred. No. 93;
 Matches   21;  Conservative    9;  Mismatches   69;  Indels    0;  Gaps    0;

Qy           3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
               ||||      : |:|         :   :::        || :| |||||      |
Db        2420 AAAAAGCACTTCATAACAAAGGAATAGCTTTAATGATGCTAAATAAAAACAAAAAAGCAA 2479

Qy          63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNU 101
               ||               || ||:                :
Db        2480 TAGAATCTTTTGAGAAAGCAATACAAATTGATAAAAATT 2518


earch completed: March  8, 2004, 07:16:15
ob time : 51 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

M protein - protein search, using sw model

un on: March 5, 2004, 09:35:33 ; Search time 55 Seconds
(without alignments)
467.488 Million cell updates/sec

itle: US-09-646-651C-1
erfect score: 467
equence: 1 TKLEDHLEGIINIFHQYSVR.....EFVVLVTDVLITAHDNIHKE 91

coring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

earched: 1586107 seqs, 282547505 residues

otal number of hits satisfying chosen parameters: 1586107

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

esult No.	Score	% Query Match	Length	DB	ID	Description
1	467	100.0	91	2	AAW01826	Aaw01826 Component
2	467	100.0	91	2	AAW93819	Aaw93819 Angiotrop
3	371	79.4	92	2	AAW03563	Aaw03563 Calcium b
4	366	78.4	90	3	AAY90765	Aay90765 Bovine CA
5	366	78.4	90	3	AAY90764	Aay90764 Bovine co
6	332	71.1	91	4	AAB31909	Aab31909 Amino aci
7	332	71.1	92	2	AAW03564	Aaw03564 Calcium b
8	332	71.1	92	2	AAW24137	Aaw24137 Human che
9	332	71.1	92	3	AAB45542	Aab45542 Human S10
10	332	71.1	92	4	AAB31911	Aab31911 Amino aci
11	332	71.1	92	4	AAB31907	Aab31907 Amino aci
12	332	71.1	92	4	AAB31908	Aab31908 Amino aci
13	332	71.1	92	7	ADA93649	Ada93649 Human cal
14	306	65.5	95	4	ABG27582	Abg27582 Novel hum
15	213.5	45.7	114	2	AAW17062	Aaw17062 Human mul
16	213.5	45.7	114	2	AAW60178	Aaw60178 Human cal
17	213.5	45.7	114	2	AAY48615	Aay48615 Human bre
18	213.5	45.7	114	3	AAY87637	Aay87637 Human cal
19	213.5	45.7	114	3	AAB45539	Aab45539 Human S10
20	213.5	45.7	114	4	ABB44613	Abb44613 Human wou
21	213.5	45.7	114	4	AAB31905	Aab31905 Amino aci
22	213.5	45.7	114	6	ABB82712	Abb82712 Human MRP
23	213.5	45.7	114	7	ADB17567	Adb17567 Human mye
24	213.5	45.7	114	7	ADE57110	Ade57110 Human Pro
25	213.5	45.7	114	7	ADE34548	Ade34548 Human mig
26	213.5	45.7	152	4	AAM39994	Aam39994 Human pol
27	206	44.1	50	3	AAY90763	Aay90763 Human EN-
28	204.5	43.8	112	7	ADB79921	Adb79921 Rat intra
29	204.5	43.8	112	7	ADE57108	Ade57108 Rat Prote
30	204.5	43.8	113	6	ABU63335	Abu63335 Rat intra
31	203	43.5	115	4	AAB31930	Aab31930 Amino aci
32	192	41.1	46	4	ABB43183	Abb43183 Peptide #
33	192	41.1	46	4	AAM37021	Aam37021 Peptide #
34	192	41.1	46	4	ABB26281	Abb26281 Protein #
35	192	41.1	46	4	AAM76914	Aam76914 Human bon
36	192	41.1	46	4	AAM64093	Aam64093 Human bra
37	192	41.1	46	4	ABG58579	Abg58579 Human liv
38	192	41.1	46	5	ABG46027	Abg46027 Human pep
39	181	38.8	95	3	AAB45544	Aab45544 Human S10
40	181	38.8	95	3	AAB45545	Aab45545 Human S10
41	181	38.8	95	5	ABG77189	Abg77189 Prostate
42	181	38.8	95	6	ADA10980	Ada10980 Human cDN
43	181	38.8	113	4	ABB12007	Abb12007 Human Ca-
44	179	38.3	181	4	AAU31075	Aau31075 Novel hum
45	178.5	38.2	113	4	ABB44612	Abb44612 Mouse wou

ALIGNMENTS

RESULT 1
AAW01826
ID AAW01826 standard; protein; 91 AA.
XX
AC AAW01826;
XX
DT 16-OCT-1997 (first entry)
XX
DE Component of bioactive metal RNA polypeptide.
XX
KW Bioactive; metal; RNA polypeptide; RNP; modulation; analysis;
KW angiogenesis; vascular state; mammalian tissue; transfer; cell;
KW genetic information; selective; alteration; nucleic acid content;
KW leukocyte; pig; monocyto-CuRNP.
XX
OS Sus scrofa.
XX
PN DE19628895-A1.
XX
PD 23-JAN-1997.
XX
PF 17-JUL-1996; 96DE-01028895.
XX
PR 17-JUL-1995; 95DE-01025992.
PR 18-AUG-1995; 95DE-01030500.
XX
PA (FRAU) FRAUNHOFER GES FOERDERUNG ANGEWANDTEN.
XX
PI Wissler JH, Logemann E, Kiesewetter S, Heilmeyer LMG;
XX
DR WPI; 1997-088586/09.
DR N-PSDB; AAT62569.
XX
PT Bioactive metal RNA polypeptide - useful for modulating angiogenesis,
PT etc.
XX
PS Claim 1; Page 15; 16pp; German.
XX
CC A novel bioactive metal RNA polypeptide (RNP) has a RNA component
CC including the sequence AAT62568 and a polypeptide component having the
CC sequence AAW01826, which is encoded by AAT62569. The RNP, or anti-RNP
CC immunoglobulins, can be used to modulate and/or analyse angiogenesis and
CC the vascular state of mammalian tissue, transfer genetic information in
CC cells and selectively alter the nucleic acid content of cells. Leukocytes
CC from pig's blood were cultured in medium, and the supernatant treated
CC with NH4 sulphate at 35, 45 and 90% saturation to precipitate protein
CC fractions. The residual supernatant was diltued to 45% NH4 sulphate
CC saturation and concentrated by ultrafiltration using a 0.5 kD membrane.

```
CC   The retenate was purified to give 8 mg of product described as monocyto-
CC   CuRNP
XX
SQ   Sequence 91 AA;

  Query Match             100.0%;  Score 467;  DB 2;  Length 91;
  Best Local Similarity   100.0%;  Pred. No. 7.5e-46;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              |||||||||||||||||||||||||||||||
Db         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91


RESULT 2
AAW93819
ID   AAW93819 standard; peptide; 91 AA.
XX
AC   AAW93819;
XX
DT   21-JUN-1999  (first entry)
XX
DE   Angiotropin related protein derived peptide.
XX
KW   Angiotropin related protein; ARP; ternary complex; S100 protein;
KW   copper-containing ribonucleoprotein; copper; cell selective;
KW   morphogenic action; blood capillary endothelial cell; confluent;
KW   non-mitogenic induction; cell phenotype; three-dimensional organoid;
KW   spatiotemporal supracellular organisation; chemotropic; blood vessel;
KW   tissue neovascularisation; angiogenesis modulation.
XX
OS   Synthetic.
XX
PN   DE19811047-C1.
XX
PD   15-APR-1999.
XX
PF   13-MAR-1998;   98DE-01011047.
XX
PR   13-MAR-1998;   98DE-01011047.
XX
PA   (FRAU ) FRAUNHOFER GES FOERDERUNG ANGEWANDTEN.
XX
PI   Kiesewetter S,  Kuhn E,  Koch-Pelster B,  Brunner H;
XX
DR   WPI; 1999-216114/19.
XX
PT   Copper-containing ribonucleoproteins - useful for modulating
PT   angiogenesis.
XX
PS   Disclosure; Page 2; 16pp; German.
XX
CC   This invention describes novel copper-containing ribonucleoproteins which
CC   are ternary complexes of an S100 protein, copper ions, and RNA comprising
CC   the following consensus sequence or its complement GGAAAAUNNNNNUNAUAUGN1-
CC   6CUNNNUUUNNNNNNAAAAN0-1UANAAACAUN0-5CUUNAGN0- 13AGAA-AUN0-16UUAGCAG where
CC   N = G, A, U or C. The ribonucleoproteins are stated to have the following
CC   properties (1) cell-selective morphogenic action in vitro on isolated
CC   primary and/or cloned blood capillary endothelial cells in culture for
CC   the non-mitogenic induction of the change in cell phenotype from the
CC   confluent state, for non-mitogenic alteration of the spatiotemporal
CC   supracellular organisation of cells into three-dimensional organoid,
CC   capillary-like structures in culture, (2) a specific chemotropic action
CC   on blood vessels in vivo, (3) induction of directional growth of blood
CC   vessels in vivo and (4) induction of neovascularisation of tissues
CC   through directed ingrowth of blood vessels. Their use for modulating
CC   angiogenesis is claimed
XX
SQ   Sequence 91 AA;

  Query Match             100.0%;  Score 467;  DB 2;  Length 91;
  Best Local Similarity   100.0%;  Pred. No. 7.5e-46;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              |||||||||||||||||||||||||||||||
Db         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91


RESULT 3
AAW03563
ID   AAW03563 standard; protein; 92 AA.
XX
AC   AAW03563;
XX
DT   01-MAY-1997  (first entry)
XX
DE   Calcium binding protein CAAF1.
XX
KW   Calcium binding protein; bovine; amniotic fluid; S100 protein family;
KW   intracellular signal transduction; squamous epithelial cell; neutrophil;
KW   macrophage; cancer; cancerous lesion; inflammation; neoplasia; cervix;
KW   squamous cell carcinoma; skin; oesophagus; CAAF1; lung; blood disease.
XX
OS   Bos taurus.
XX
PN   EP731166-A2.
XX
PD   11-SEP-1996.
XX
PF   04-DEC-1995;   95EP-00119045.
XX
PR   06-MAR-1995;   95JP-00045564.
PR   06-MAR-1995;   95JP-00070468.
XX
PA   (TOFU ) TONEN CORP.
PA   (HITO/) HITOMI J.
XX
PI   Hitomi J,  Yamaguchi K,  Yamamura T,  Kimura T;
XX
DR   WPI; 1996-403989/41.
DR   N-PSDB; AAT39345.
XX
PT   New human or bovine calcium binding protein and related nucleic acid - is
PT   a marker for inflammation, neoplasia, skin and blood diseases.
XX
PS   Claim 1; Page 21; 36pp; English.
XX
CC   This sequence represents the CAAF1 calcium-binding protein isolated from
CC   bovine amniotic fluid. CAAF1 belongs to the S100 protein family, which
CC   includes calcyclin, MRP8, and MRP14. Intracellular calcium ion
CC   concentration is one of the key factors for intracellular signal
CC   transduction. The calcium signals are transduced by various calcium-
CC   binding proteins, such as the protein encoded by this sequence. CAAF1 is
CC   normally expressed in squamous epithelial cells, neutrophils and
CC   macrophages, but atypical epithelial cells are negative for CAAF1 and
CC   overexpression is observed in several types of cancer cells and
CC   neutrophils/macrophages infiltrating cancerous lesions. Detection of
CC   CAAF1 (using antibodies in usual immunoassays) can be used to diagnose
CC   (or monitor) inflammation, neoplasia (particularly squamous cell
CC   carcinoma of the skin, oesophagus, lung and cervix), and skin and blood
CC   diseases
XX
SQ   Sequence 92 AA;

  Query Match             79.4%;  Score 371;  DB 2;  Length 92;
  Best Local Similarity   81.3%;  Pred. No. 8.2e-35;
  Matches   74;  Conservative    9;  Mismatches    8;  Indels    0;  Gaps    0;
```

Qy 1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
Db 2 TKLEDHLEGIINIFHQYSVRVGHFDTLNKRELKQLITKELPKTLQNTKDQPTIDKIFQDL 61

Qy 61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
Db 62 DADKDGAVSFEEFVVLVSRVLKTAHIDIHKE 92

RESULT 4
AAY90765
ID AAY90765 standard; protein; 90 AA.
XX
AC AAY90765;
XX
DT 18-AUG-2000 (first entry)
XX
DE Bovine CAAF1 acid sequence SEQ ID NO:4.
XX
KW Bovine; EN-RAGE; extracellular novel RAGE binding protein;
KW receptor for advanced glycation endproduct; inflammation; inhibition;
KW antiinflammatory; immunoglobulin; cell surface molecule; septic shock;
KW systemic lupus erythematosus; inflammatory lupus nephritis; endotoxaemia;
KW autoimmune disorder; inflammatory disorder.
XX
OS Bos taurus.
XX
PN WO200020621-A1.
XX
PD 13-APR-2000.
XX
PF 06-OCT-1999; 99WO-US023303.
XX
PR 06-OCT-1998; 98US-00167705.
PR 05-MAR-1999; 99US-00263312.
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Schmidt AM, Stern D;
XX
DR WPI; 2000-303794/26.
XX
PT New human EN-RAGE (extracellular novel receptor for advanced glycation
PT end products) peptide, useful for identifying anti-inflammatory compounds
PT that inhibit its interaction with RAGE.
XX
PS Claim 2; Page 41; 132pp; English.
XX
CC The present invention describes an isolated human EN-RAGE (extracellular
CC novel receptor for advanced glycation end products) peptide (P1). The EN-
CC RAGE peptide binds to RAGE which is a member of the immunoglobulin
CC superfamily of cell-surface molecules. A compound capable of inhibiting
CC the interaction of EN-RAGE with RAGE is useful for the suppression of
CC inflammation resulting from systemic lupus erythematosus, inflammatory
CC lupus nephritis, septic shock, endotoxaemia, or an autoimmune or
CC inflammatory disorder in which the recruitment of EN-RAGE containing
CC inflammatory cells occurs. The compound is also useful for the treatment
CC of systemic lupus erythematosus, inflammatory lupus nephritis in a
CC subject. The human EN-RAGE peptide is useful for identifying compounds
CC that inhibit its interaction with RAGE. The present sequence represents
CC the bovine CAAF1 which shows homology to the human EN-RAGE N-terminal
CC amino acid sequence
XX
SQ Sequence 90 AA;

Query Match 78.4%; Score 366; DB 3; Length 90;
Best Local Similarity 81.1%; Pred. No. 3e-34;
Matches 73; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

Qy 1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
Db 1 TKLEDHLEGIINIFHQYSVRVGHFDTLNKRELKQLITKELPKTLQNTKDQPTIDKIFQDL 60

Qy 61 DANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
Db 61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90

RESULT 5
AAY90764
ID AAY90764 standard; protein; 90 AA.
XX
AC AAY90764;
XX
DT 18-AUG-2000 (first entry)
XX
DE Bovine corneal antigen (B-COAg) acid sequence SEQ ID NO:3.
XX
KW Bovine; EN-RAGE; extracellular novel RAGE binding protein;
KW receptor for advanced glycation endproduct; inflammation; inhibition;
KW antiinflammatory; immunoglobulin; cell surface molecule; septic shock;
KW systemic lupus erythematosus; inflammatory lupus nephritis; endotoxaemia;
KW autoimmune disorder; inflammatory disorder.
XX
OS Bos taurus.
XX
PN WO200020621-A1.
XX
PD 13-APR-2000.
XX
PF 06-OCT-1999; 99WO-US023303.
XX
PR 06-OCT-1998; 98US-00167705.
PR 05-MAR-1999; 99US-00263312.
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Schmidt AM, Stern D;
XX
DR WPI; 2000-303794/26.
XX
PT New human EN-RAGE (extracellular novel receptor for advanced glycation
PT end products) peptide, useful for identifying anti-inflammatory compounds
PT that inhibit its interaction with RAGE.
XX
PS Claim 2; Page 41; 132pp; English.
XX
CC The present invention describes an isolated human EN-RAGE (extracellular
CC novel receptor for advanced glycation end products) peptide (P1). The EN-
CC RAGE peptide binds to RAGE which is a member of the immunoglobulin
CC superfamily of cell-surface molecules. A compound capable of inhibiting
CC the interaction of EN-RAGE with RAGE is useful for the suppression of
CC inflammation resulting from systemic lupus erythematosus, inflammatory
CC lupus nephritis, septic shock, endotoxaemia, or an autoimmune or
CC inflammatory disorder in which the recruitment of EN-RAGE containing
CC inflammatory cells occurs. The compound is also useful for the treatment
CC of systemic lupus erythematosus, inflammatory lupus nephritis in a
CC subject. The human EN-RAGE peptide is useful for identifying compounds
CC that inhibit its interaction with RAGE. The present sequence represents
CC the bovine corneal antigen which shows homology to the human EN-RAGE N-
CC terminal amino acid sequence
XX
SQ Sequence 90 AA;

Query Match 78.4%; Score 366; DB 3; Length 90;
Best Local Similarity 81.1%; Pred. No. 3e-34;
Matches 73; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

Qy 1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
Db 1 TKLEDHLEGIINIFHQYSVRVGHFDTLNKRELKQLITKELPKTLQNTKDQPTIDKIFQDL 60

Qy 61 DANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
Db 61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90

```
RESULT 6
AAB31909
ID   AAB31909 standard; protein; 91 AA.
XX
AC   AAB31909;
XX
DT   15-MAY-2001  (first entry)
XX
DE   Amino acid sequence of a human protein.
XX
KW   Human; perlecan; retinol-binding plasma protein; calgranulin B; vaccine;
KW   ganglioside GM2 activator; saposin B; degenerative disease; glial cell;
KW   neurological disease; auto-immune disease; multiple sclerosis; toxicity;
KW   Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
KW   rheumatoid polyarthritis; lupus erythematosus; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200105422-A2.
XX
PD   25-JAN-2001.
XX
PF   17-JUL-2000; 2000WO-FR002057.
XX
PR   15-JUL-1999;   99FR-00009372.
XX
PA   (INMR ) BIOMERIEUX STELHYS.
XX
PI   Roecklin D,  Kolbe H,  Charles M,  Malcus C,  Santoro L,  Perron H;
XX
DR   WPI; 2001-159475/16.
XX
PT   Detecting, preventing and treating degenerative, neurological and
PT   autoimmune diseases, particularly multiple sclerosis, using specified
PT   polypeptides or related nucleic acid or ligand.
XX
PS   Claim 1; Page 167; 209pp; French.
XX
CC   The present sequence represents a human protein, which is used in the
CC   method of the invention. The specification describes a method which uses
CC   at least one polypeptide or polynucleotide sequence belonging to the
CC   perlecan, precursor of the retinol-binding plasma protein, precursor of
CC   the ganglioside GM2 activator, calgranulin B or saposin B protein
CC   families. The method is used for detecting, preventing or treating a
CC   degenerative, neurological and/or auto-immune disease. The
CC   polynucleotides and polypeptides are used for diagnosis, prognosis,
CC   prevention and treatment of multiple sclerosis (in its various forms and
CC   phases). They may also be useful in cases of e.g. Alzheimer's and
CC   Parkinson's diseases, amyotrophic lateral sclerosis, rheumatoid
CC   polyarthritis and lupus erythematosus, including use as vaccines and in
CC   gene therapy (expression of sense or antisense sequences). They can also
CC   be used to assess efficacy of potential therapeutic agents, particularly
CC   compounds that reduce or inhibit toxicity towards glial cells
XX
SQ   Sequence 91 AA;

  Query Match             71.1%;  Score 332;  DB 4;  Length 91;
  Best Local Similarity   70.3%;  Pred. No. 2.5e-30;
  Matches   64;  Conservative   10;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ||||:|||||:|||||||||| ||:||| | |||||:|||| ||:|| ||:  ||:||| |
Db          1 TKLEEHLEGIVNIFHQYSVRKGHFDTLSKGELKQLLTKELANTIKNIKDKAVIDEIFQGL 60

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              |||||||| |:||: ||   |  || : |||
Db         61 DANQDEQVDFQEFISLVAIALKAAHYHTHKE 91


RESULT 7
AAW03564
ID   AAW03564 standard; protein; 92 AA.
XX
AC   AAW03564;
XX
DT   01-MAY-1997  (first entry)
XX
DE   Calcium binding protein CAAF1.
XX
KW   Calcium binding protein; human; amniotic fluid; S100 protein family;
KW   intracellular signal transduction; squamous epithelial cell; neutrophil;
KW   macrophage; cancer; cancerous lesion; inflammation; neoplasia; cervix;
KW   squamous cell carcinoma; skin; oesophagus; CAAF1; lung; blood disease.
XX
OS   Homo sapiens.
XX
PN   EP731166-A2.
XX
PD   11-SEP-1996.
XX
PF   04-DEC-1995;   95EP-00119045.
XX
PR   06-MAR-1995;   95JP-00045564.
PR   06-MAR-1995;   95JP-00070468.
XX
PA   (TOFU ) TONEN CORP.
PA   (HITO/) HITOMI J.
XX
PI   Hitomi J,  Yamaguchi K,  Yamamura T,  Kimura T;
XX
DR   WPI; 1996-403989/41.
DR   N-PSDB; AAT39346.
XX
PT   New human or bovine calcium binding protein and related nucleic acid - is
PT   a marker for inflammation, neoplasia, skin and blood diseases.
XX
PS   Claim 1; Page 24; 36pp; English.
XX
CC   This sequence represents the CAAF1 calcium-binding protein isolated from
CC   human amniotic fluid. CAAF1 belongs to the S100 protein family, which
CC   includes calcyclin, MRP8, and MRP14. Intracellular calcium ion
CC   concentration is one of the key factors for intracellular signal
CC   transduction. The calcium signals are transduced by various calcium-
CC   binding proteins, such as this protein. CAAF1 is normally expressed in
CC   squamous epithelial cells, neutrophils and macrophages, but atypical
CC   epithelial cells are negative for CAAF1 and overexpression is observed in
CC   several types of cancer cells and neutrophils/macrophages infiltrating
CC   cancerous lesions. Detection of CAAF1 (using antibodies in usual
CC   immunoassays) can be used to diagnose (or monitor) inflammation,
CC   neoplasia (particularly squamous cell carcinoma of the skin, oesophagus,
CC   lung and cervix), and skin and blood diseases
XX
SQ   Sequence 92 AA;

  Query Match             71.1%;  Score 332;  DB 2;  Length 92;
  Best Local Similarity   70.3%;  Pred. No. 2.5e-30;
  Matches   64;  Conservative   10;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ||||:|||||:|||||||||| ||:||| | |||||:|||| ||:|| ||:  ||:||| |
Db          2 TKLEEHLEGIVNIFHQYSVRKGHFDTLSKGELKQLLTKELANTIKNIKDKAVIDEIFQGL 61

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              |||||||| |:||: ||   |  || : |||
Db         62 DANQDEQVDFQEFISLVAIALKAAHYHTHKE 92


RESULT 8
AAW24137
ID   AAW24137 standard; protein; 92 AA.
XX
AC   AAW24137;
XX
DT   28-JAN-1998  (first entry)
```

XX
DE Human chemotactic cytokine I.
XX
KW chemotactic cytokine; tumour; autoimmune disease; antagonist; agonist.
XX
OS Homo sapiens.
XX
PN WO9723640-A1.
XX
PD 03-JUL-1997.
XX
PF 26-DEC-1995; 95WO-US016871.
XX
PR 26-DEC-1995; 95WO-US016871.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ni J, Yu G, Alfonso P, Gentz R, Su JY;
XX
DR WPI; 1997-351075/32.
DR N-PSDB; AAT85774.
XX
PT DNA encoding chemotactic cytokine I - used to treat, e.g. tumours,
PT chronic infection, leukaemia, etc.
XX
PS Claim 12; Page 48-49; 64pp; English.
XX
CC This is a human chemotactic cytokine I polypeptide. The encoding
CC polynucleotide, along with a vector and a host cell can be used for the
CC recombinant production of the chemotactic cytokine. Cytokine agonists and
CC antagonists can be used for the treatment of a patient requiring a
CC chemotactic cytokine I and for the treatment of a patient requiring the
CC inhibition of a chemotactic cytokine I polypeptide, respectively. The
CC chemotactic cytokine is used to treat tumours, chronic infection,
CC leukaemia and T-cell mediated autoimmune diseases
XX
SQ Sequence 92 AA;

Query Match 71.1%; Score 332; DB 2; Length 92;
Best Local Similarity 70.3%; Pred. No. 2.5e-30;
Matches 64; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

Qy 1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
Db 2 TKLEEHLEGIVNIFHQYSVRKGHFDTLSKGELKQLLTKELANTIKNIKDKAVIDEIFQGL 61

Qy 61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
Db 62 DANQDEQVDFQEFISLVAIALKAAHYHTHKE 92

RESULT 9
AAB45542
ID AAB45542 standard; protein; 92 AA.
XX
AC AAB45542;
XX
DT 22-FEB-2001 (first entry)
XX
DE Human S100A12 protein.
XX
KW S100 protein; human; treatment; cardiomyopathy; cardiac insufficiency;
KW calcium-binding protein; calcium homeostasis; cardiac muscle;
KW pumping capacity; myocardial cell; systolic calcium ion release;
KW sarcoplasmic reticulum; cardiac disease; hypertension; rhythm disorder;
KW valve defect.
XX
OS Homo sapiens.
XX
PN DE19915485-A1.
XX
PD 19-OCT-2000.
XX
PF 07-APR-1999; 99DE-01015485.
XX
PR 07-APR-1999; 99DE-01015485.
XX
PA (KATU/) KATUS H A.
PA (REMP/) REMPPIS A.
XX
PI Katus HA, Remppis A;
XX
DR WPI; 2000-673510/66.
DR N-PSDB; AAC81812.
XX
PT Composition containing S100 protein, corresponding nucleic acid or
PT vector, useful for treating cardiomyopathy and cardiac insufficiency.
XX
PS Claim 35; Page 20; 36pp; German.
XX
CC This invention describes a novel composition for treating primary or
CC secondary cardiomyopathy or cardiac insufficiency contains at least one
CC S100 protein (I) or nucleic acid (II) encoding (I), or their mutants or
CC fragments, or a gene transfer vector containing (II), optionally
CC formulated with auxiliaries and/or carriers. (I) are calcium-binding
CC proteins involved in calcium homeostasis, so their overexpression in
CC cardiac muscle will improve pumping capacity (and overall capacity) of
CC the heart. In cultured myocardial cells they increase the contraction and
CC relaxation rates associated with increased systolic calcium ion release
CC from the sarcoplasmic reticulum (SR) and calcium re-uptake by SR. (I) are
CC used to treat cardiomyopathy (CMP) where inherited or caused by
CC spontaneous mutations and ischemic CMP caused by arteriosclerosis,
CC dilative CMP caused by toxic/infectious disease, cardiac disease caused
CC by pulmonary and/or arterial hypertension, and structural disease caused
CC by rhythm disorders or valve defects, generally any condition associated
CC with reduced contractile force. Unlike calmodulin, which is expressed
CC ubiquitously, (I) show tissue-specific expression and treat the
CC underlying defect in the sarcoplasmic reticulum (SR) that causes cardiac
CC disease
XX
SQ Sequence 92 AA;

Query Match 71.1%; Score 332; DB 3; Length 92;
Best Local Similarity 70.3%; Pred. No. 2.5e-30;
Matches 64; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

Qy 1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
Db 2 TKLEEHLEGIVNIFHQYSVRKGHFDTLSKGELKQLLTKELANTIKNIKDKAVIDEIFQGL 61

Qy 61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
Db 62 DANQDEQVDFQEFISLVAIALKAAHYHTHKE 92

RESULT 10
AAB31911
ID AAB31911 standard; protein; 92 AA.
XX
AC AAB31911;
XX
DT 15-MAY-2001 (first entry)
XX
DE Amino acid sequence of a human protein.
XX
KW Human; perlecan; retinol-binding plasma protein; calgranulin B; vaccine;
KW ganglioside GM2 activator; saposin B; degenerative disease; glial cell;
KW neurological disease; auto-immune disease; multiple sclerosis; toxicity;
KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
KW rheumatoid polyarthritis; lupus erythematosus; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200105422-A2.
XX
PD 25-JAN-2001.

XX
PF 17-JUL-2000; 2000WO-FR002057.
XX
PR 15-JUL-1999; 99FR-00009372.
XX
PA (INMR) BIOMERIEUX STELHYS.
XX
PI Roecklin D, Kolbe H, Charles M, Malcus C, Santoro L, Perron H;
XX
DR WPI; 2001-159475/16.
XX
PT Detecting, preventing and treating degenerative, neurological and
PT autoimmune diseases, particularly multiple sclerosis, using specified
PT polypeptides or related nucleic acid or ligand.
XX
PS Claim 1; Page 168; 209pp; French.
XX
CC The present sequence represents a human protein, which is used in the
CC method of the invention. The specification describes a method which uses
CC at least one polypeptide or polynucleotide sequence belonging to the
CC perlecan, precursor of the retinol-binding plasma protein, precursor of
CC the ganglioside GM2 activator, calgranulin B or saposin B protein
CC families. The method is used for detecting, preventing or treating a
CC degenerative, neurological and/or auto-immune disease. The
CC polynucleotides and polypeptides are used for diagnosis, prognosis,
CC prevention and treatment of multiple sclerosis (in its various forms and
CC phases). They may also be useful in cases of e.g. Alzheimer's and
CC Parkinson's diseases, amyotrophic lateral sclerosis, rheumatoid
CC polyarthritis and lupus erythematosus, including use as vaccines and in
CC gene therapy (expression of sense or antisense sequences). They can also
CC be used to assess efficacy of potential therapeutic agents, particularly
CC compounds that reduce or inhibit toxicity towards glial cells
XX
SQ Sequence 92 AA;

Query Match 71.1%; Score 332; DB 4; Length 92;
Best Local Similarity 70.3%; Pred. No. 2.5e-30;
Matches 64; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

```
Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ||||:|||||:|||||||||| ||:||| | |||||:|||| ||:|| ||:  ||:||| |
Db          2 TKLEEHLEGIVNIFHQYSVRKGHFDTLSKGELKQLLTKELANTIKNIKDKAVIDEIFQGL 61

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              ||||||| |:||: ||   |  || : |||
Db         62 DANQDEQVDFQEFISLVAIALKAAHYHTHKE 92
```

RESULT 11
AAB31907
ID AAB31907 standard; protein; 92 AA.
XX
AC AAB31907;
XX
DT 15-MAY-2001 (first entry)
XX
DE Amino acid sequence of a human protein.
XX
KW Human; perlecan; retinol-binding plasma protein; calgranulin B; vaccine;
KW ganglioside GM2 activator; saposin B; degenerative disease; glial cell;
KW neurological disease; auto-immune disease; multiple sclerosis; toxicity;
KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
KW rheumatoid polyarthritis; lupus erythematosus; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200105422-A2.
XX
PD 25-JAN-2001.
XX
PF 17-JUL-2000; 2000WO-FR002057.
XX
PR 15-JUL-1999; 99FR-00009372.
XX
PA (INMR) BIOMERIEUX STELHYS.
XX
PI Roecklin D, Kolbe H, Charles M, Malcus C, Santoro L, Perron H;
XX
DR WPI; 2001-159475/16.
XX
PT Detecting, preventing and treating degenerative, neurological and
PT autoimmune diseases, particularly multiple sclerosis, using specified
PT polypeptides or related nucleic acid or ligand.
XX
PS Claim 1; Page 166-167; 209pp; French.
XX
CC The present sequence represents a human protein, which is used in the
CC method of the invention. The specification describes a method which uses
CC at least one polypeptide or polynucleotide sequence belonging to the
CC perlecan, precursor of the retinol-binding plasma protein, precursor of
CC the ganglioside GM2 activator, calgranulin B or saposin B protein
CC families. The method is used for detecting, preventing or treating a
CC degenerative, neurological and/or auto-immune disease. The
CC polynucleotides and polypeptides are used for diagnosis, prognosis,
CC prevention and treatment of multiple sclerosis (in its various forms and
CC phases). They may also be useful in cases of e.g. Alzheimer's and
CC Parkinson's diseases, amyotrophic lateral sclerosis, rheumatoid
CC polyarthritis and lupus erythematosus, including use as vaccines and in
CC gene therapy (expression of sense or antisense sequences). They can also
CC be used to assess efficacy of potential therapeutic agents, particularly
CC compounds that reduce or inhibit toxicity towards glial cells
XX
SQ Sequence 92 AA;

Query Match 71.1%; Score 332; DB 4; Length 92;
Best Local Similarity 70.3%; Pred. No. 2.5e-30;
Matches 64; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

```
Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ||||:|||||:|||||||||| ||:||| | |||||:|||| ||:|| ||:  ||:||| |
Db          2 TKLEEHLEGIVNIFHQYSVRKGHFDTLSKGELKQLLTKELANTIKNIKDKAVIDEIFQGL 61

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              ||||||| |:||: ||   |  || : |||
Db         62 DANQDEQVDFQEFISLVAIALKAAHYHTHKE 92
```

RESULT 12
AAB31908
ID AAB31908 standard; protein; 92 AA.
XX
AC AAB31908;
XX
DT 15-MAY-2001 (first entry)
XX
DE Amino acid sequence of a human protein.
XX
KW Human; perlecan; retinol-binding plasma protein; calgranulin B; vaccine;
KW ganglioside GM2 activator; saposin B; degenerative disease; glial cell;
KW neurological disease; auto-immune disease; multiple sclerosis; toxicity;
KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
KW rheumatoid polyarthritis; lupus erythematosus; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200105422-A2.
XX
PD 25-JAN-2001.
XX
PF 17-JUL-2000; 2000WO-FR002057.
XX
PR 15-JUL-1999; 99FR-00009372.
XX
PA (INMR) BIOMERIEUX STELHYS.
XX
PI Roecklin D, Kolbe H, Charles M, Malcus C, Santoro L, Perron H;

XX
RR WPI; 2001-159475/16.
XX
PT Detecting, preventing and treating degenerative, neurological and
PT autoimmune diseases, particularly multiple sclerosis, using specified
PT polypeptides or related nucleic acid or ligand.
XX
PS Claim 1; Page 167; 209pp; French.
XX
CC The present sequence represents a human protein, which is used in the
CC method of the invention. The specification describes a method which uses
CC at least one polypeptide or polynucleotide sequence belonging to the
CC perlecan, precursor of the retinol-binding plasma protein, precursor of
CC the ganglioside GM2 activator, calgranulin B or saposin B protein
CC families. The method is used for detecting, preventing or treating a
CC degenerative, neurological and/or auto-immune disease. The
CC polynucleotides and polypeptides are used for diagnosis, prognosis,
CC prevention and treatment of multiple sclerosis (in its various forms and
CC phases). They may also be useful in cases of e.g. Alzheimer's and
CC Parkinson's diseases, amyotrophic lateral sclerosis, rheumatoid
CC polyarthritis and lupus erythematosus, including use as vaccines and in
CC gene therapy (expression of sense or antisense sequences). They can also
CC be used to assess efficacy of potential therapeutic agents, particularly
CC compounds that reduce or inhibit toxicity towards glial cells
XX
SQ Sequence 92 AA;

Query Match 71.1%; Score 332; DB 4; Length 92;
Best Local Similarity 70.3%; Pred. No. 2.5e-30;
Matches 64; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

Qy 1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
Db 2 TKLEEHLEGIVNIFHQYSVRKGHFDTLSKGELKQLLTKELANTIKNIKDKAVIDEIFQGL 61

Qy 61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
Db 62 DANQDEQVDFQEFISLVAIALKAAHYHTHKE 92

RESULT 13
DA93649
DE ADA93649 standard; protein; 92 AA.
XX
AC ADA93649;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human calgranulin C protein SEQ ID NO:2.
XX
KW inflammatory disease; calgranulin C; antiinflammatory; gene therapy;
KW vasculitis; Kawasaki disease; cystic fibrosis;
KW chronic inflammatory disease; ulcerative colitis; Crohn's disease;
KW chronic bronchitis; inflammatory arthritis; psoriatic arthritis;
KW rheumatoid arthritis; seronegative arthritis;
KW systemic onset juvenile rheumatoid arthritis; SOJRA; Still's disease;
KW acute inflammation; human.
XX
OS Homo sapiens.
XX
PN WO2003069341-A2.
XX
PD 21-AUG-2003.
XX
PF 17-FEB-2003; 2003WO-EP001575.
XX
PR 15-FEB-2002; 2002US-00077600.
XX
PA (SWIT-) SWITCH BIOTECH AG.
PA (SORG/) SORG C.
PA (ROTH/) ROTH J.
XX
PI Sorg C, Roth J;
XX
DR WPI; 2003-671681/63.
DR N-PSDB; ADA93648.
XX
PT Diagnosing, treating or preventing inflammatory diseases comprises
PT determining the amount and/or concentration of CALGRANULIN C polypeptide
PT and/or nucleic acids encoding the polypeptide present in a biological
PT sample.
XX
PS Claim 7; Page 64; 64pp; English.
XX
CC The present invention describes a method for diagnosing inflammatory
CC diseases, which comprises determining the amount and/or concentration of
CC calgranulin C polypeptide and/or nucleic acids encoding the polypeptide
CC present in the biological sample. Also described are methods for treating
CC or preventing an inflammatory disease in a mammal, and medical treatment
CC of the mammal, where the treatment is based on the stage of the disease
CC to be treated or prevented. Calgranulin C has antiinflammatory activity
CC and can be used in gene therapy. The method is useful for diagnosing,
CC treating or preventing inflammatory diseases, e.g. vasculitis
CC (particularly Kawasaki disease), cystic fibrosis, chronic inflammatory
CC diseases like ulcerative colitis or Crohn's disease, chronic bronchitis,
CC inflammatory arthritis (e.g. psoriatic arthritis, rheumatoid arthritis or
CC seronegative arthritis), systemic onset juvenile rheumatoid arthritis
CC (SOJRA or Still's disease), acute inflammation above the background of a
CC chronic inflammation, an acquired infection on the background of an
CC inflammatory disease, or an exacerbation of an already present disease.
CC The method is also useful for diagnosing specific stages of inflammatory
CC diseases, for determining the risk of relapse, and for discriminating
CC between diseases with similar symptoms. The present sequence represents
CC human calgranulin C, which is used in the exemplification of the present
CC invention.
XX
SQ Sequence 92 AA;

Query Match 71.1%; Score 332; DB 7; Length 92;
Best Local Similarity 70.3%; Pred. No. 2.5e-30;
Matches 64; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

Qy 1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
Db 2 TKLEEHLEGIVNIFHQYSVRKGHFDTLSKGELKQLLTKELANTIKNIKDKAVIDEIFQGL 61

Qy 61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
Db 62 DANQDEQVDFQEFISLVAIALKAAHYHTHKE 92

RESULT 14
ABG27582
ID ABG27582 standard; protein; 95 AA.
XX
AC ABG27582;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #27573.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.

```
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS91769.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 20; SEQ ID NO 57941; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC   amino acid sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 95 AA;

  Query Match             65.5%;  Score 306;  DB 4;  Length 95;
  Best Local Similarity   67.0%;  Pred. No. 2.5e-27;
  Matches   61;  Conservative   10;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ||||:|||||:||  | ||| ||:||| | |||||:|||| ||:|| ||:  ||:||| |
Db          5 TKLEEHLEGIVNISPQXSVRKGHFDTLSKGELKQLLTKELANTIKNIKDKAVIDEIFQGL 64

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              |||||||| |:||: ||   |  || : |||
Db         65 DANQDEQVDFQEFISLVAIALKAAHYHTHKE 95


RESULT 15
AAW17062
ID   AAW17062 standard; protein; 114 AA.
XX
AC   AAW17062;
XX
DT   16-JUL-1997  (first entry)
XX
DE   Human multidrug resistance protein 14 (MRP14).
XX
KW   Genetic engineering; MRP; multidrug resistance protein; transgenic;
KW   animal model; cell death inhibition; apoptosis; cell proliferation; HIV;
KW   human immunodeficiency virus; cancer; cystic fibrosis; neoplasia; tumour.
XX
OS   Homo sapiens.
XX
PN   US5614397-A.
XX
PD   25-MAR-1997.
XX
PF   22-FEB-1994;   94US-00200016.
XX
PR   22-FEB-1994;   94US-00200016.
XX
PA   (STRD ) UNIV LELAND STANFORD JUNIOR.
XX
PI   Weissman I,  Lagasse E;
XX
DR   WPI; 1997-224943/20.
DR   N-PSDB; AAT68322.
XX
PT   Increasing life-span of mammalian haemato-lymphoid cells by transforming
PT   stem cells - with construct contg. cell-specific transcription initiator
PT   and gene encoding protein that increases lifetime, useful for drug
PT   screening and treatment.
XX
PS   Example 1; Col 33-34; 34pp; English.
XX
CC   AAW17062 is the human MRP14 protein. The transcriptional initiator of the
CC   MRP8 gene was used in a construct for expressing an open reading frame
CC   that increase the lifespan of a mammalian haematolymphoid cell, e.g. the
CC   mammalian bcl-2 gene, a CFTR (cystic fibrosis transmembrane regulator)
CC   gene, the herpes virus thymidine kinase gene or an oncogene.
CC   Haematolymphoid cells are especially neutrophils and the construct
CC   doubles the lifespan of transgenic cells. Transgenic cells or transgenic
CC   animals produced are used for screening for substances and treatments
CC   that prevent or promote cell death. They can also be returned to the
CC   patient to modulate apoptosis, i.e. in the treatment of disorders related
CC   to abnormal cell proliferation or death. Typical applications are
CC   treatment of viral diseases, including HIV; cancer and cystic fibrosis
XX
SQ   Sequence 114 AA;

  Query Match             45.7%;  Score 213.5;  DB 2;  Length 114;
  Best Local Similarity   45.1%;  Pred. No. 1.3e-16;
  Matches   41;  Conservative   25;  Mismatches   24;  Indels    1;  Gaps    1;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTL-KNTKDQGTIDKIFQN 59
              ::|| ::| ||| ||||||:||| ||| : | |:|: |:| | | |  |:: |: | ::
Db          6 SQLERNIETIINTFHQYSVKLGHPDTLNQGEFKELVRKDLQNFLKKENKNEKVIEHIMED 65

Qy         60 LDANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
              || | |:|:||:||::|:  :   :|: :|:
Db         66 LDTNADKQLSFEEFIMLMARLTWASHEKMHE 96


Search completed: March  5, 2004, 09:46:44
Job time : 56 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

M protein - protein search, using sw model

un on: March 5, 2004, 09:45:44 ; Search time 23 Seconds
(without alignments)
204.259 Million cell updates/sec

itle: US-09-646-651C-1
erfect score: 467
equence: 1 TKLEDHLEGIINIFHQYSVR.....EFVVLVTDVLITAHDNIHKE 91

coring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

earched: 389414 seqs, 51625971 residues

otal number of hits satisfying chosen parameters: 389414

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

esult No.	Score	% Query Match	Length	DB	ID	Description
1	467	100.0	91	3	US-08-794-000-2	Sequence 2, Appli
2	371	79.4	92	2	US-08-568-310D-19	Sequence 19, Appl
3	371	79.4	92	4	US-09-270-455-19	Sequence 19, Appl
4	366	78.4	90	4	US-09-263-312-3	Sequence 3, Appli
5	366	78.4	90	4	US-09-826-589-3	Sequence 3, Appli
6	366	78.4	90	4	US-09-826-589-4	Sequence 4, Appli
7	332	71.1	92	2	US-08-568-310D-20	Sequence 20, Appl
8	332	71.1	92	4	US-09-270-455-20	Sequence 20, Appl
9	235	50.3	51	2	US-08-568-310D-2	Sequence 2, Appli
10	235	50.3	51	4	US-09-270-455-2	Sequence 2, Appli
11	213.5	45.7	114	1	US-08-385-241-3	Sequence 3, Appli
12	213.5	45.7	114	4	US-09-214-272-4	Sequence 4, Appli
13	211.5	45.3	109	1	US-07-987-272A-8	Sequence 8, Appli
14	206	44.1	50	4	US-09-263-312-2	Sequence 2, Appli
15	206	44.1	50	4	US-09-826-589-2	Sequence 2, Appli
16	204.5	43.8	113	2	US-08-918-727-7	Sequence 7, Appli
17	204.5	43.8	113	3	US-09-205-680A-7	Sequence 7, Appli
18	181	38.8	95	4	US-09-919-172-102	Sequence 102, App
19	181	38.8	95	4	US-09-976-594-467	Sequence 467, App
20	164	35.1	92	2	US-09-051-589-1	Sequence 1, Appli
21	163	34.9	88	1	US-07-987-272A-1	Sequence 1, Appli
22	163	34.9	89	1	US-07-987-272A-14	Sequence 14, Appl
23	162	34.7	91	1	US-07-987-272A-11	Sequence 11, Appl
24	162	34.7	92	2	US-08-918-727-5	Sequence 5, Appli
25	162	34.7	92	3	US-09-205-680A-5	Sequence 5, Appli
26	156	33.4	101	1	US-08-190-560-2	Sequence 2, Appli
27	156	33.4	101	1	US-08-469-277-2	Sequence 2, Appli
28	156	33.4	101	2	US-08-468-946-2	Sequence 2, Appli
29	156	33.4	101	2	US-08-468-942-2	Sequence 2, Appli
30	156	33.4	101	4	US-09-298-625-2	Sequence 2, Appli
31	151	32.3	93	1	US-07-987-272A-7	Sequence 7, Appli
32	151	32.3	93	1	US-07-987-272A-16	Sequence 16, Appl
33	151	32.3	93	1	US-08-385-241-1	Sequence 1, Appli
34	151	32.3	93	4	US-09-214-272-2	Sequence 2, Appli
35	150	32.1	105	2	US-08-918-727-6	Sequence 6, Appli
36	150	32.1	105	3	US-09-205-680A-6	Sequence 6, Appli
37	148	31.7	97	1	US-07-662-198B-2	Sequence 2, Appli
38	140.5	30.1	98	2	US-08-918-727-3	Sequence 3, Appli
39	140.5	30.1	98	3	US-09-205-680A-3	Sequence 3, Appli
40	140.5	30.1	98	3	US-09-048-889-11	Sequence 11, Appl
41	140	30.0	92	4	US-09-621-976-7537	Sequence 7537, Ap
42	137	29.3	76	1	US-07-987-272A-17	Sequence 17, Appl
43	136	29.1	88	4	US-09-621-976-7524	Sequence 7524, Ap
44	136	29.1	88	4	US-09-621-976-7526	Sequence 7526, Ap
45	136	29.1	89	1	US-07-987-272A-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-08-794-000-2
; Sequence 2, Application US/08794000
; Patent No. 6087123
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Metal-Containing Ribonucleotide Polypeptides
; NUMBER OF SEQUENCES: 4
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/794,000
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/DE96/01337
; FILING DATE: 17-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 25 992.0
; FILING DATE: 17-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 30 500.0
; FILING DATE: 18-AUG-1995
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 91 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-794-000-2

Query Match 100.0%; Score 467; DB 3; Length 91;
Best Local Similarity 100.0%; Pred. No. 4.4e-49;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              |||||||||||||||||||||||||||||||
Db         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
```

RESULT 2
US-08-568-310D-19
; Sequence 19, Application US/08568310D

```
; Patent No. 5976832
;  GENERAL INFORMATION:
;    APPLICANT:  HITOMI, JIRO
;    APPLICANT:  YAMAGUCHI, KEN
;    APPLICANT:  YAMAMURA,TOKUJIRO
;    APPLICANT:  KIMURA, TATSUJI
;    TITLE OF INVENTION:  NOVEL CALCIUM-BINDING PROTEINS
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  WYATT,GERBER,MELLER & O'ROURKE
;      STREET:  99 PARK AVENUE
;      STREET:  6th FLOOR
;      CITY:  NEW YORK CITY
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10016
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE,3.50 INCH,720 Kb
;      MEDIUM TYPE:  STORAGE
;      COMPUTER:  IBM-PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS 6.2
;      SOFTWARE:  WORDPERFECT 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/568,310D
;      FILING DATE:  DECEMBER 6, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  7-70468 and 7-45564(both Japan)
;      FILING DATE:  3/6/95 and 3/6/95,respectively
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KLEIN, MILTON
;      REGISTRATION NUMBER:  27101
;      REFERENCE/DOCKET NUMBER:  3316
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)953-3350
;      TELEFAX:  (212)953-3352
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  92
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    PUBLICATION INFORMATION:
;      RELEVANT RESIDUES IN SEQ ID NO:  19:
;      RELEVANT RESIDUES IN SEQ ID NO:  FROM 1 TO 92
US-08-568-310D-19

  Query Match             79.4%;  Score 371;  DB 2;  Length 92;
  Best Local Similarity   81.3%;  Pred. No. 1.7e-37;
  Matches   74;  Conservative    9;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ||||||||||||||||||||:||:||| |||||||||||| ||:||||| |||||||:|
Db          2 TKLEDHLEGIINIFHQYSVRVGHFDTLNKRELKQLITKELPKTLQNTKDQPTIDKIFQDL 61

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              ||::|  |||:|||||: || ||| :||||
Db         62 DADKDGAVSFEEFVVLVSRVLKTAHIDIHKE 92


RESULT 3
US-09-270-455-19
; Sequence 19, Application US/09270455
; Patent No. 6313267
;  GENERAL INFORMATION:
;    APPLICANT:  HITOMI, JIRO
;    APPLICANT:  YAMAGUCHI, KEN
;    APPLICANT:  YAMAMURA,TOKUJIRO
;    APPLICANT:  KIMURA, TATSUJI
;    TITLE OF INVENTION:  NOVEL CALCIUM-BINDING PROTEINS
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  WYATT,GERBER,MELLER & O'ROURKE
;      STREET:  99 PARK AVENUE
;      STREET:  6th FLOOR
;      CITY:  NEW YORK CITY
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10016
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE,3.50 INCH,720 Kb
;      MEDIUM TYPE:  STORAGE
;      COMPUTER:  IBM-PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS 6.2
;      SOFTWARE:  WORDPERFECT 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/270,455
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/568,310
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KLEIN, MILTON
;      REGISTRATION NUMBER:  27101
;      REFERENCE/DOCKET NUMBER:  3316
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)953-3350
;      TELEFAX:  (212)953-3352
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  92
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    PUBLICATION INFORMATION:
;      RELEVANT RESIDUES IN SEQ ID NO:  19: FROM 1 TO 92
US-09-270-455-19

  Query Match             79.4%;  Score 371;  DB 4;  Length 92;
  Best Local Similarity   81.3%;  Pred. No. 1.7e-37;
  Matches   74;  Conservative    9;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ||||||||||||||||||||:||:||| |||||||||||| ||:||||| |||||||:|
Db          2 TKLEDHLEGIINIFHQYSVRVGHFDTLNKRELKQLITKELPKTLQNTKDQPTIDKIFQDL 61

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              ||::|  |||:|||||: || ||| :||||
Db         62 DADKDGAVSFEEFVVLVSRVLKTAHIDIHKE 92


RESULT 4
US-09-263-312-3
; Sequence 3, Application US/09263312
; Patent No. 6555340
; GENERAL INFORMATION:
;  APPLICANT:  Schmidt, Ann Marie
;  APPLICANT:  Stern, David
;  TITLE OF INVENTION:  Extracellular No. 6555340el RAGE Binding Protein (EN-RAGE) and
;  TITLE OF INVENTION:  Uses Thereof
;  FILE REFERENCE:  0575/55873-A
;  CURRENT APPLICATION NUMBER: US/09/263,312
;  CURRENT FILING DATE:  1999-03-05
;  NUMBER OF SEQ ID NOS:  5
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 90
;   TYPE: PRT
;   ORGANISM: Human
US-09-263-312-3

  Query Match             78.4%;  Score 366;  DB 4;  Length 90;
  Best Local Similarity   81.1%;  Pred. No. 6.7e-37;
```

Matches 73; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

```
y          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
             |||||||||||||||||||:||:||| |||||||||||| ||:||||| |||||||:|
b          1 TKLEDHLEGIINIFHQYSVRVGHFDTLNKRELKQLITKELPKTLQNTKDQPTIDKIFQDL 60

y         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
             ||::|  |||:||||||: || ||| :|||
b         61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90
```

ESULT 5
S-09-826-589-3
Sequence 3, Application US/09826589
Patent No. 6670136
GENERAL INFORMATION:
APPLICANT: Schmidt, Ann Marie
APPLICANT: Stern, David
TITLE OF INVENTION: EXTRACELLULAR NOVEL RAGE BINDING PROTEIN (EN-RAGE) AND USES THERE
FILE REFERENCE: 0575/55873-B-PCT-US
CURRENT APPLICATION NUMBER: US/09/826,589
CURRENT FILING DATE: 2001-04-05
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 90
TYPE: PRT
ORGANISM: Bovine
S-09-826-589-3

Query Match 78.4%; Score 366; DB 4; Length 90;
Best Local Similarity 81.1%; Pred. No. 6.7e-37;
Matches 73; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

```
y          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
             |||||||||||||||||||:||:||| |||||||||||| ||:||||| |||||||:|
b          1 TKLEDHLEGIINIFHQYSVRVGHFDTLNKRELKQLITKELPKTLQNTKDQPTIDKIFQDL 60

y         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
             ||::|  |||:||||||: || ||| :|||
b         61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90
```

ESULT 6
S-09-826-589-4
Sequence 4, Application US/09826589
Patent No. 6670136
GENERAL INFORMATION:
APPLICANT: Schmidt, Ann Marie
APPLICANT: Stern, David
TITLE OF INVENTION: EXTRACELLULAR NOVEL RAGE BINDING PROTEIN (EN-RAGE) AND USES THERE
FILE REFERENCE: 0575/55873-B-PCT-US
CURRENT APPLICATION NUMBER: US/09/826,589
CURRENT FILING DATE: 2001-04-05
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 90
TYPE: PRT
ORGANISM: Bovine
S-09-826-589-4

Query Match 78.4%; Score 366; DB 4; Length 90;
Best Local Similarity 81.1%; Pred. No. 6.7e-37;
Matches 73; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

```
y          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
             |||||||||||||||||||:||:||| |||||||||||| ||:||||| |||||||:|
o          1 TKLEDHLEGIINIFHQYSVRVGHFDTLNKRELKQLITKELPKTLQNTKDQPTIDKIFQDL 60

y         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
             ||::|  |||:||||||: || ||| :|||
o         61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90
```

RESULT 7
US-08-568-310D-20
; Sequence 20, Application US/08568310D
; Patent No. 5976832
; GENERAL INFORMATION:
; APPLICANT: HITOMI, JIRO
; APPLICANT: YAMAGUCHI, KEN
; APPLICANT: YAMAMURA,TOKUJIRO
; APPLICANT: KIMURA, TATSUJI
; TITLE OF INVENTION: NOVEL CALCIUM-BINDING PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WYATT,GERBER,MELLER & O'ROURKE
; STREET: 99 PARK AVENUE
; STREET: 6th FLOOR
; CITY: NEW YORK CITY
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE,3.50 INCH,720 Kb
; MEDIUM TYPE: STORAGE
; COMPUTER: IBM-PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS 6.2
; SOFTWARE: WORDPERFECT 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,310D
; FILING DATE: DECEMBER 6, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 7-70468 and 7-45564(both Japan)
; FILING DATE: 3/6/95 and 3/6/95,respectively
; ATTORNEY/AGENT INFORMATION:
; NAME: KLEIN, MILTON
; REGISTRATION NUMBER: 27101
; REFERENCE/DOCKET NUMBER: 3316
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)953-3350
; TELEFAX: (212)953-3352
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 92
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; PUBLICATION INFORMATION:
; RELEVANT RESIDUES IN SEQ ID NO: 20:
; RELEVANT RESIDUES IN SEQ ID NO: FROM 1 TO 92
US-08-568-310D-20

Query Match 71.1%; Score 332; DB 2; Length 92;
Best Local Similarity 70.3%; Pred. No. 8.7e-33;
Matches 64; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

```
Qy         1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
             ||||:|||||:|||||||||| ||:||| | |||||:|||| ||:|| ||:  ||:||| |
Db         2 TKLEEHLEGIVNIFHQYSVRKGHFDTLSKGELKQLLTKELANTIKNIKDKAVIDEIFQGL 61

Qy        61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
             |||||||| |:||: ||   |  || : |||
Db        62 DANQDEQVDFQEFISLVAIALKAAHYHTHKE 92
```

RESULT 8
US-09-270-455-20
; Sequence 20, Application US/09270455
; Patent No. 6313267
; GENERAL INFORMATION:
; APPLICANT: HITOMI, JIRO
; APPLICANT: YAMAGUCHI, KEN

```
;     APPLICANT:  YAMAMURA,TOKUJIRO
;     APPLICANT:  KIMURA, TATSUJI
;     TITLE OF INVENTION:  NOVEL CALCIUM-BINDING PROTEINS
;     NUMBER OF SEQUENCES:  20
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  WYATT,GERBER,MELLER & O'ROURKE
;       STREET:  99 PARK AVENUE
;       STREET:  6th FLOOR
;       CITY:  NEW YORK CITY
;       STATE:  NEW YORK
;       COUNTRY:  USA
;       ZIP:  10016
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  DISKETTE,3.50 INCH,720 Kb
;       MEDIUM TYPE:  STORAGE
;       COMPUTER:  IBM-PC COMPATIBLE
;       OPERATING SYSTEM:  PC-DOS 6.2
;       SOFTWARE:  WORDPERFECT 6.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/270,455
;       FILING DATE:
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/568,310
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  KLEIN, MILTON
;       REGISTRATION NUMBER:  27101
;       REFERENCE/DOCKET NUMBER:  3316
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212)953-3350
;       TELEFAX:  (212)953-3352
;   INFORMATION FOR SEQ ID NO:  20:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  92
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     PUBLICATION INFORMATION:
;       RELEVANT RESIDUES IN SEQ ID NO:  20: FROM 1 TO 92
US-09-270-455-20

  Query Match             71.1%;  Score 332;  DB 4;  Length 92;
  Best Local Similarity   70.3%;  Pred. No. 8.7e-33;
  Matches   64;  Conservative   10;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ||||:|||||:|||||||| ||:||| | |||||:|||| ||:|| ||:  ||:||| |
Db          2 TKLEEHLEGIVNIFHQYSVRKGHFDTLSKGELKQLLTKELANTIKNIKDKAVIDEIFQGL 61

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              |||||||| |:|| ||   |  || : |||
Db         62 DANQDEQVDFQEFISLVAIALKAAHYHTHKE 92


RESULT 9
US-08-568-310D-2
; Sequence 2, Application US/08568310D
; Patent No. 5976832
;  GENERAL INFORMATION:
;    APPLICANT:  HITOMI, JIRO
;    APPLICANT:  YAMAGUCHI, KEN
;    APPLICANT:  YAMAMURA,TOKUJIRO
;    APPLICANT:  KIMURA, TATSUJI
;    TITLE OF INVENTION:  NOVEL CALCIUM-BINDING PROTEINS
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  WYATT,GERBER,MELLER & O'ROURKE
;      STREET:  99 PARK AVENUE
;      STREET:  6th FLOOR
;      CITY:  NEW YORK CITY
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10016
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE,3.50 INCH,720 Kb
;      MEDIUM TYPE:  STORAGE
;      COMPUTER:  IBM-PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS 6.2
;      SOFTWARE:  WORDPERFECT 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/568,310D
;      FILING DATE:  DECEMBER 6, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  7-70468 and 7-45564(both Japan)
;      FILING DATE:  3/6/95 and 3/6/95,respectively
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KLEIN, MILTON
;      REGISTRATION NUMBER:  27101
;      REFERENCE/DOCKET NUMBER:  3316
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)953-3350
;      TELEFAX:  (212)953-3352
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  51
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    PUBLICATION INFORMATION:
;      RELEVANT RESIDUES IN SEQ ID NO:  2:
;      RELEVANT RESIDUES IN SEQ ID NO:  FROM 1 TO 51
US-08-568-310D-2

  Query Match             50.3%;  Score 235;  DB 2;  Length 51;
  Best Local Similarity   90.0%;  Pred. No. 2e-21;
  Matches   45;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQ 50
              |||||||||||||||||||:||:||| ||||||||||||| ||:|||||
Db          1 TKLEDHLEGIINIFHQYSVRVGHFDTLNKRELKQLITKELPKTLQNTKDQ 50


RESULT 10
US-09-270-455-2
; Sequence 2, Application US/09270455
; Patent No. 6313267
;  GENERAL INFORMATION:
;    APPLICANT:  HITOMI, JIRO
;    APPLICANT:  YAMAGUCHI, KEN
;    APPLICANT:  YAMAMURA,TOKUJIRO
;    APPLICANT:  KIMURA, TATSUJI
;    TITLE OF INVENTION:  NOVEL CALCIUM-BINDING PROTEINS
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  WYATT,GERBER,MELLER & O'ROURKE
;      STREET:  99 PARK AVENUE
;      STREET:  6th FLOOR
;      CITY:  NEW YORK CITY
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10016
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE,3.50 INCH,720 Kb
;      MEDIUM TYPE:  STORAGE
;      COMPUTER:  IBM-PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS 6.2
;      SOFTWARE:  WORDPERFECT 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/270,455
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/568,310
```

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: KLEIN, MILTON
REGISTRATION NUMBER: 27101
REFERENCE/DOCKET NUMBER: 3316
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)953-3350
TELEFAX: (212)953-3352
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 51
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
PUBLICATION INFORMATION:
RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 51
JS-09-270-455-2

Query Match 50.3%; Score 235; DB 4; Length 51;
Best Local Similarity 90.0%; Pred. No. 2e-21;
Matches 45; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQ 50
Db 1 TKLEDHLEGIINIFHQYSVRVGHFDTLNKRELKQLITKELPKTLQNTKDQ 50

RESULT 11
JS-08-385-241-3
Sequence 3, Application US/08385241
Patent No. 5776348
GENERAL INFORMATION:
APPLICANT: Selengut Ph.D., Jeremy D.
APPLICANT: Orme-Johnson Ph.D., William H.
APPLICANT: Dretler M.D., Stephen P.
APPLICANT: Asakura M.D., Hirotaka
TITLE OF INVENTION: SYSTEM AND METHOD FOR INHIBITING
TITLE OF INVENTION: FORMATION OF CRYSTALLINE STRUCTURES THAT INCLUDE STRUVITE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Choate, Hall & Stewart
STREET: 53 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2891
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/385,241
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Herschbach Ph.D., Brenda M.
REGISTRATION NUMBER: P-39,223
REFERENCE/DOCKET NUMBER: 492611-000 (MIT6915)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-5175
TELEFAX: (617) 248-4000
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 114 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: hMRP-14 protein

US-08-385-241-3

Query Match 45.7%; Score 213.5; DB 1; Length 114;
Best Local Similarity 45.1%; Pred. No. 2.3e-18;
Matches 41; Conservative 25; Mismatches 24; Indels 1; Gaps 1;

Qy 1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTL-KNTKDQGTIDKIFQN 59
Db 6 SQLERNIETIINTFHQYSVKLGHPDTLNQGEFKELVRKDLQNFLKKENKNEKVIEHIMED 65

Qy 60 LDANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
Db 66 LDTNADKQLSFEEFIMLMARLTWASHEKMHE 96

RESULT 12
US-09-214-272-4
; Sequence 4, Application US/09214272
; Patent No. 6620790
; GENERAL INFORMATION:
; APPLICANT: Siegenthaler, Georges
; TITLE OF INVENTION: Isolated Arachidonic Acid-Binding Heteromer and its Use in
; TITLE OF INVENTION: Cosmetics and Pharmaceutics
; FILE REFERENCE: 016800-254
; CURRENT APPLICATION NUMBER: US/09/214,272
; CURRENT FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: PCT/FR97/01164
; PRIOR FILING DATE: 1997-06-30
; PRIOR APPLICATION NUMBER: FR 96/08219
; PRIOR FILING DATE: 1996-07-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of MRP-14.
US-09-214-272-4

Query Match 45.7%; Score 213.5; DB 4; Length 114;
Best Local Similarity 45.1%; Pred. No. 2.3e-18;
Matches 41; Conservative 25; Mismatches 24; Indels 1; Gaps 1;

Qy 1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTL-KNTKDQGTIDKIFQN 59
Db 6 SQLERNIETIINTFHQYSVKLGHPDTLNQGEFKELVRKDLQNFLKKENKNEKVIEHIMED 65

Qy 60 LDANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
Db 66 LDTNADKQLSFEEFIMLMARLTWASHEKMHE 96

RESULT 13
US-07-987-272A-8
; Sequence 8, Application US/07987272A
; Patent No. 5731166
; GENERAL INFORMATION:
; APPLICANT: Geczy, C., Simpson, R. J. and Lackmann, M
; TITLE OF INVENTION: No. 5731166el Chemotactic Factor
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cushman Darby & Cushman
; STREET: 1100 New York Avenue, N. W., Ninth Floor, East Tower
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/987,272A
;       FILING DATE:  05-MAR-1993
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  AU PK 2127
;       FILING DATE:  05-FEB-1990
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  AU PK 4463
;       FILING DATE:  05-SEP-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Brinkman, David W
;       REGISTRATION NUMBER:  20,817
;       REFERENCE/DOCKET NUMBER:  DWB/1925/200259
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-861 3000
;       TELEFAX:  202-822 0944
;       TELEX:  6714627 CUSH
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  109 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-07-987-272A-8

  Query Match             45.3%;  Score 211.5;  DB 1;  Length 109;
  Best Local Similarity   45.1%;  Pred. No. 3.7e-18;
  Matches   41;  Conservative   25;  Mismatches   24;  Indels    1;  Gaps    1;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTL-KNTKDQGTIDKIFQN 59
              ::|| ::| ||| |||||:||| ||| : | |:|: |:| | | |  |::  |: | ::
Db          1 SQLERNIETIINTFHQYSVKLGHPDTLNQGEFKELVRKDLQNFLKKENKNELVIEHIMED 60

Qy         60 LDANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
              || | |:|:||:||::|:  :   :|: :|:
Db         61 LDTNADKQLSFEEFIMLMARLTWASHEKMHE 91


RESULT 14
US-09-263-312-2
; Sequence 2, Application US/09263312
; Patent No. 6555340
; GENERAL INFORMATION:
;  APPLICANT:  Schmidt, Ann Marie
;  APPLICANT:  Stern, David
;  TITLE OF INVENTION:  Extracellular No. 6555340el RAGE Binding Protein (EN-RAGE) and
;  TITLE OF INVENTION:  Uses Thereof
;  FILE REFERENCE:  0575/55873-A
;  CURRENT APPLICATION NUMBER: US/09/263,312
;  CURRENT FILING DATE:  1999-03-05
;  NUMBER OF SEQ ID NOS:  5
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 50
;   TYPE: PRT
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (47)
;   OTHER INFORMATION: Xaa at this position is unknown
US-09-263-312-2

  Query Match             44.1%;  Score 206;  DB 4;  Length 50;
  Best Local Similarity   82.0%;  Pred. No. 6.3e-18;
  Matches   41;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQ 50
              |||||||||||||| ||||||:||:||| | ||||| ||||| ||:| |||
Db          1 TKLEDHLEGIINIGHQYSVRVGHFDTLNKYELKQLGTKELPKTLQNXKDQ 50


RESULT 15
US-09-826-589-2
; Sequence 2, Application US/09826589
; Patent No. 6670136
; GENERAL INFORMATION:
;  APPLICANT: Schmidt, Ann Marie
;  APPLICANT:  Stern, David
;  TITLE OF INVENTION: EXTRACELLULAR NOVEL RAGE BINDING PROTEIN (EN-RAGE) AND USES THERI
;  FILE REFERENCE: 0575/55873-B-PCT-US
;  CURRENT APPLICATION NUMBER: US/09/826,589
;  CURRENT FILING DATE:  2001-04-05
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 50
;   TYPE: PRT
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (47)..(47)
;   OTHER INFORMATION: x=any amino acid
US-09-826-589-2

  Query Match             44.1%;  Score 206;  DB 4;  Length 50;
  Best Local Similarity   82.0%;  Pred. No. 6.3e-18;
  Matches   41;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQ 50
              |||||||||||||| ||||||:||:||| | ||||| ||||| ||:| |||
Db          1 TKLEDHLEGIINIGHQYSVRVGHFDTLNKYELKQLGTKELPKTLQNXKDQ 50


Search completed: March  5, 2004, 09:50:52
Job time : 23 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

)M protein - protein search, using sw model

!un on: March 5, 2004, 09:46:49 ; Search time 34 Seconds
(without alignments)
565.145 Million cell updates/sec

!itle: US-09-646-651C-1
)erfect score: 467
!equence: 1 TKLEDHLEGIINIFHQYSVR.....EFVVLVTDVLITAHDNIHKE 91

!coring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

!earched: 809742 seqs, 211153259 residues

)otal number of hits satisfying chosen parameters: 809742

!inimum DB seq length: 0
!aximum DB seq length: 2000000000

)ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

)atabase : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

esult No.	Score	% Query Match	Length	DB	ID	Description
1	366	78.4	90	9	US-09-826-589-3	Sequence 3, Appli
2	366	78.4	90	9	US-09-826-589-4	Sequence 4, Appli
3	366	78.4	90	9	US-09-872-185B-11	Sequence 11, Appl
4	366	78.4	90	9	US-09-872-185B-12	Sequence 12, Appl
5	332	71.1	92	14	US-10-077-600-2	Sequence 2, Appli
6	213.5	45.7	114	14	US-10-134-841-4	Sequence 4, Appli
7	213.5	45.7	114	14	US-10-308-279-32	Sequence 32, Appl
8	213.5	45.7	114	15	US-10-116-275-225	Sequence 225, App
9	213.5	45.7	114	15	US-10-131-410-146	Sequence 146, App
10	206	44.1	50	9	US-09-826-589-2	Sequence 2, Appli
11	206	44.1	50	9	US-09-872-185B-9	Sequence 9, Appli
12	204.5	43.8	112	14	US-10-205-219-161	Sequence 161, App
13	204.5	43.8	113	10	US-09-492-026-7	Sequence 7, Appli
14	192	41.1	46	9	US-09-864-761-41579	Sequence 41579, A
15	181	38.8	95	9	US-09-919-172-102	Sequence 102, App
16	181	38.8	95	9	US-09-981-353-98	Sequence 98, Appl
17	178.5	38.2	113	14	US-10-134-841-3	Sequence 3, Appli
18	171	36.6	91	14	US-10-106-698-6907	Sequence 6907, Ap
19	163	34.9	89	14	US-10-134-841-1	Sequence 1, Appli
20	163	34.9	101	9	US-09-393-433-2	Sequence 2, Appli
21	163	34.9	101	9	US-09-781-509-2	Sequence 2, Appli
22	163	34.9	101	14	US-10-269-643-2	Sequence 2, Appli
23	162	34.7	92	10	US-09-492-026-5	Sequence 5, Appli
24	162	34.7	92	10	US-09-919-039-184	Sequence 184, App
25	160	34.3	94	14	US-10-097-340-270	Sequence 270, App
26	159	34.0	89	14	US-10-316-253-46	Sequence 46, Appl
27	156	33.4	101	9	US-09-393-433-1	Sequence 1, Appli
28	156	33.4	101	9	US-09-781-509-1	Sequence 1, Appli
29	156	33.4	101	13	US-10-067-618-2	Sequence 2, Appli
30	156	33.4	101	13	US-10-135-152-2	Sequence 2, Appli
31	156	33.4	101	14	US-10-269-643-1	Sequence 1, Appli
32	154	33.0	93	15	US-10-094-886-58	Sequence 58, Appl
33	151	32.3	93	14	US-10-134-841-2	Sequence 2, Appli
34	151	32.3	477	15	US-10-161-927-62	Sequence 62, Appl
35	150	32.1	105	10	US-09-492-026-6	Sequence 6, Appli
36	150	32.1	105	14	US-10-097-340-272	Sequence 272, App
37	150	32.1	105	14	US-10-301-822-177	Sequence 177, App
38	150	32.1	134	9	US-09-925-302-694	Sequence 694, App
39	143	30.6	97	14	US-10-097-340-274	Sequence 274, App
40	143	30.6	97	14	US-10-171-311-206	Sequence 206, App
41	143	30.6	97	15	US-10-236-031B-40	Sequence 40, Appl
42	141	30.2	84	10	US-09-849-138-34	Sequence 34, Appl
43	140.5	30.1	98	10	US-09-492-026-3	Sequence 3, Appli
44	140.5	30.1	110	15	US-10-264-049-2820	Sequence 2820, Ap
45	138	29.6	186	15	US-10-264-049-4104	Sequence 4104, Ap

ALIGNMENTS

RESULT 1
US-09-826-589-3
; Sequence 3, Application US/09826589
; Patent No. US20020106726A1
; GENERAL INFORMATION:
; APPLICANT: Schmidt, Ann Marie
; APPLICANT: Stern, David
; TITLE OF INVENTION: EXTRACELLULAR NOVEL RAGE BINDING PROTEIN (EN-RAGE) AND USES THERE
; FILE REFERENCE: 0575/55873-B-PCT-US
; CURRENT APPLICATION NUMBER: US/09/826,589
; CURRENT FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Bovine
US-09-826-589-3

Query Match 78.4%; Score 366; DB 9; Length 90;
Best Local Similarity 81.1%; Pred. No. 1.2e-32;
Matches 73; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

```
Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              |||||||||||||||||||:||:||| |||||||||||| ||:||||| |||||||:|
Db          1 TKLEDHLEGIINIFHQYSVRVGHFDTLNKRELKQLITKELPKTLQNTKDQPTIDKIFQDL 60

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
              ||::|  |||:||||||: || ||| :|||
Db         61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90
```

RESULT 2
US-09-826-589-4
; Sequence 4, Application US/09826589
; Patent No. US20020106726A1
; GENERAL INFORMATION:
; APPLICANT: Schmidt, Ann Marie

```
;  APPLICANT:  Stern, David
;  TITLE OF INVENTION: EXTRACELLULAR NOVEL RAGE BINDING PROTEIN (EN-RAGE) AND USES THERE
;  FILE REFERENCE: 0575/55873-B-PCT-US
;  CURRENT APPLICATION NUMBER: US/09/826,589
;  CURRENT FILING DATE:  2001-04-05
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 90
;   TYPE: PRT
;   ORGANISM: Bovine
US-09-826-589-4

  Query Match             78.4%;  Score 366;  DB 9;  Length 90;
  Best Local Similarity   81.1%;  Pred. No. 1.2e-32;
  Matches   73;  Conservative    9;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              |||||||||||||||||||||:||:||| |||||||||||| ||:||||| |||||||:|
Db          1 TKLEDHLEGIINIFHQYSVRVGHFDTLNKRELKQLITKELPKTLQNTKDQPTIDKIFQDL 60

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
              ||::|  |||:||||||: || ||| :|||
Db         61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90


RESULT 3
US-09-872-185B-11
; Sequence 11, Application US/09872185B
; Patent No. US20020122799A1
; GENERAL INFORMATION:
;  APPLICANT: Stern, David M.
;  APPLICANT:  Herold, Kevan
;  APPLICANT:  Yan, Shi Du
;  APPLICANT:  Schmidt, Ann Marie
;  APPLICANT:  Lamster, Ira
;  TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION
;  FILE REFERENCE: 0575/64080
;  CURRENT APPLICATION NUMBER: US/09/872,185B
;  CURRENT FILING DATE:  2001-06-01
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
;   LENGTH: 90
;   TYPE: PRT
;   ORGANISM: Bovine
US-09-872-185B-11

  Query Match             78.4%;  Score 366;  DB 9;  Length 90;
  Best Local Similarity   81.1%;  Pred. No. 1.2e-32;
  Matches   73;  Conservative    9;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              |||||||||||||||||||:||:||| |||||||||||| ||:||||| |||||||:|
Db          1 TKLEDHLEGIINIFHQYSVRVGHFDTLNKRELKQLITKELPKTLQNTKDQPTIDKIFQDL 60

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
              ||::|  |||:||||||: || ||| :|||
Db         61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90


RESULT 4
US-09-872-185B-12
; Sequence 12, Application US/09872185B
; Patent No. US20020122799A1
; GENERAL INFORMATION:
;  APPLICANT: Stern, David M.
;  APPLICANT:  Herold, Kevan
;  APPLICANT:  Yan, Shi Du
;  APPLICANT:  Schmidt, Ann Marie
;  APPLICANT:  Lamster, Ira
;  TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION
;  FILE REFERENCE: 0575/64080
;  CURRENT APPLICATION NUMBER: US/09/872,185B
;  CURRENT FILING DATE:  2001-06-01
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
;   LENGTH: 90
;   TYPE: PRT
;   ORGANISM: Bovine
US-09-872-185B-12

  Query Match             78.4%;  Score 366;  DB 9;  Length 90;
  Best Local Similarity   81.1%;  Pred. No. 1.2e-32;
  Matches   73;  Conservative    9;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              |||||||||||||||||||:||:||| |||||||||||| ||:||||| |||||||:|
Db          1 TKLEDHLEGIINIFHQYSVRVGHFDTLNKRELKQLITKELPKTLQNTKDQPTIDKIFQDL 60

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
              ||::|  |||:||||||: || ||| :|||
Db         61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90


RESULT 5
US-10-077-600-2
; Sequence 2, Application US/10077600
; Publication No. US20030175713A1
; GENERAL INFORMATION:
;  APPLICANT: Switch Biotech AG
;  TITLE OF INVENTION: Method for diagnosis of inflammatory diseases using Calgranulin (
;  FILE REFERENCE: S30274US
;  CURRENT APPLICATION NUMBER: US/10/077,600
;  CURRENT FILING DATE:  2002-06-14
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 92
;   TYPE: PRT
;   ORGANISM: homo sapiens
US-10-077-600-2

  Query Match             71.1%;  Score 332;  DB 14;  Length 92;
  Best Local Similarity   70.3%;  Pred. No. 6.6e-29;
  Matches   64;  Conservative   10;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ||||:|||||:||||||||| ||:||| | |||||:|||| ||:|| ||:  ||:||| |
Db          2 TKLEEHLEGIVNIFHQYSVRKGHFDTLSKGELKQLLTKELANTIKNIKDKAVIDEIFQGL 61

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              |||||||| |:|| ||   |  || : |||
Db         62 DANQDEQVDFQEFISLVAIALKAAHYHTHKE 92


RESULT 6
US-10-134-841-4
; Sequence 4, Application US/10134841
; Publication No. US20030003482A1
; GENERAL INFORMATION:
;  APPLICANT: HALLE, JORN-PETER
;  APPLICANT:  GOPPELT, ANDREAS
;  TITLE OF INVENTION: MRP8/MRP14 heterodimer, or its
;  TITLE OF INVENTION:  individual components in combination, for treating and/or
;  TITLE OF INVENTION:  preventing skin diseases, wounds and/or wound-healing
;  TITLE OF INVENTION:  disturbances, having a reduced quantity of MRP8/MRP14
;  TITLE OF INVENTION:  heterodimers
;  FILE REFERENCE: 50125/031002
;  CURRENT APPLICATION NUMBER: US/10/134,841
;  CURRENT FILING DATE:  2002-04-29
;  PRIOR APPLICATION NUMBER: US 60/322,925
;  PRIOR FILING DATE: 2001-09-17
;  PRIOR APPLICATION NUMBER: DE 10121254.2
;  PRIOR FILING DATE: 2001-04-30
```

NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 4
LENGTH: 114
TYPE: PRT
ORGANISM: Homo sapien
S-10-134-841-4

Query Match 45.7%; Score 213.5; DB 14; Length 114;
Best Local Similarity 45.1%; Pred. No. 7.7e-16;
Matches 41; Conservative 25; Mismatches 24; Indels 1; Gaps 1;

y 1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTL-KNTKDQGTIDKIFQN 59
::|| ::| ||| |||||||:||| ||| : | |:|: |:| | | | |:: |: | ::
b 6 SQLERNIETIINTFHQYSVKLGHPDTLNQGEFKELVRKDLQNFLKKENKNEKVIEHIMED 65

y 60 LDANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
|| | |:|:||:||::|: : :|: :|:
b 66 LDTNADKQLSFEEFIMLMARLTWASHEKMHE 96

ESULT 7
S-10-308-279-32
Sequence 32, Application US/10308279
Publication No. US20030170742A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE DEVELOPMENT
TITLE OF INVENTION: RHEUMATOID ARTHRITIS
FILE REFERENCE: D0190 NP
CURRENT APPLICATION NUMBER: US/10/308,279
CURRENT FILING DATE: 2002-12-03
PRIOR APPLICATION NUMBER: 60/337,429
PRIOR FILING DATE: 2001-12-03
NUMBER OF SEQ ID NOS: 91
SOFTWARE: PatentIn version 3.1
SEQ ID NO 32
LENGTH: 114
TYPE: PRT
ORGANISM: homo sapiens
S-10-308-279-32

Query Match 45.7%; Score 213.5; DB 14; Length 114;
Best Local Similarity 45.1%; Pred. No. 7.7e-16;
Matches 41; Conservative 25; Mismatches 24; Indels 1; Gaps 1;

y 1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTL-KNTKDQGTIDKIFQN 59
::|| ::| ||| |||||||:||| ||| : | |:|: |:| | | | |:: |: | ::
b 6 SQLERNIETIINTFHQYSVKLGHPDTLNQGEFKELVRKDLQNFLKKENKNEKVIEHIMED 65

y 60 LDANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
|| | |:|:||:||::|: : :|: :|:
b 66 LDTNADKQLSFEEFIMLMARLTWASHEKMHE 96

ESULT 8
S-10-116-275-225
Sequence 225, Application US/10116275
Publication No. US20030211476A1
GENERAL INFORMATION:
APPLICANT: Elan Pharmaceutical Technology
APPLICANT: O'Mahony, Daniel J.
APPLICANT: Brayden, David
APPLICANT: Byrne, Daragh
APPLICANT: Lambkin, Imelda
APPLICANT: Higgins, Lisa
TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
TITLE OF INVENTION: Compositions Targeting Peyer's Patches and M Cell Receptors
FILE REFERENCE: E1067/20087
CURRENT APPLICATION NUMBER: US/10/116,275
CURRENT FILING DATE: 2002-10-04
NUMBER OF SEQ ID NOS: 349
SOFTWARE: PatentIn version 3.1

; SEQ ID NO 225
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-275-225

Query Match 45.7%; Score 213.5; DB 15; Length 114;
Best Local Similarity 45.1%; Pred. No. 7.7e-16;
Matches 41; Conservative 25; Mismatches 24; Indels 1; Gaps 1;

Qy 1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTL-KNTKDQGTIDKIFQN 59
::|| ::| ||| |||||||:||| ||| : | |:|: |:| | | | |:: |: | ::
Db 6 SQLERNIETIINTFHQYSVKLGHPDTLNQGEFKELVRKDLQNFLKKENKNEKVIEHIMED 65

Qy 60 LDANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
|| | |:|:||:||::|: : :|: :|:
Db 66 LDTNADKQLSFEEFIMLMARLTWASHEKMHE 96

RESULT 9
US-10-131-410-146
; Sequence 146, Application US/10131410
; Publication No. US20030235915A1
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM TISSUE OF BREAST
; TITLE OF INVENTION: TUMORS
; FILE REFERENCE: SCH-1763
; CURRENT APPLICATION NUMBER: US/10/131,410
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 09/646,673
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: PCT/DE99/00908
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 202
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 146
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-410-146

Query Match 45.7%; Score 213.5; DB 15; Length 114;
Best Local Similarity 45.1%; Pred. No. 7.7e-16;
Matches 41; Conservative 25; Mismatches 24; Indels 1; Gaps 1;

Qy 1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTL-KNTKDQGTIDKIFQN 59
::|| ::| ||| |||||||:||| ||| : | |:|: |:| | | | |:: |: | ::
Db 6 SQLERNIETIINTFHQYSVKLGHPDTLNQGEFKELVRKDLQNFLKKENKNEKVIEHIMED 65

Qy 60 LDANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
|| | |:|:||:||::|: : :|: :|:
Db 66 LDTNADKQLSFEEFIMLMARLTWASHEKMHE 96

RESULT 10
US-09-826-589-2
; Sequence 2, Application US/09826589
; Patent No. US20020106726A1
; GENERAL INFORMATION:
; APPLICANT: Schmidt, Ann Marie
; APPLICANT: Stern, David
; TITLE OF INVENTION: EXTRACELLULAR NOVEL RAGE BINDING PROTEIN (EN-RAGE) AND USES THEREO
; FILE REFERENCE: 0575/55873-B-PCT-US
; CURRENT APPLICATION NUMBER: US/09/826,589
; CURRENT FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (47)..(47)
; OTHER INFORMATION: x=any amino acid
US-09-826-589-2

Query Match 44.1%; Score 206; DB 9; Length 50;
Best Local Similarity 82.0%; Pred. No. 1.9e-15;
Matches 41; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQ 50
Db 1 TKLEDHLEGIINIGHQYSVRVGHFDTLNKYELKQLGTKELPKTLQNXKDQ 50

RESULT 11
US-09-872-185B-9
; Sequence 9, Application US/09872185B
; Patent No. US20020122799A1
; GENERAL INFORMATION:
; APPLICANT: Stern, David M.
; APPLICANT: Herold, Kevan
; APPLICANT: Yan, Shi Du
; APPLICANT: Schmidt, Ann Marie
; APPLICANT: Lamster, Ira
; TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION
; FILE REFERENCE: 0575/64080
; CURRENT APPLICATION NUMBER: US/09/872,185B
; CURRENT FILING DATE: 2001-06-01
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Bovine
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (47)..(47)
; OTHER INFORMATION: Where Xaa = unknown
US-09-872-185B-9

Query Match 44.1%; Score 206; DB 9; Length 50;
Best Local Similarity 82.0%; Pred. No. 1.9e-15;
Matches 41; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQ 50
Db 1 TKLEDHLEGIINIGHQYSVRVGHFDTLNKYELKQLGTKELPKTLQNXKDQ 50

RESULT 12
US-10-205-219-161
; Sequence 161, Application US/10205219
; Publication No. US20030138803A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pinnock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018200
; CURRENT APPLICATION NUMBER: US/10/205,219
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 161
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Intracellular calcium binding protein
US-10-205-219-161

Query Match 43.8%; Score 204.5; DB 14; Length 112;
Best Local Similarity 41.8%; Pred. No. 7.3e-15;
Matches 38; Conservative 24; Mismatches 28; Indels 1; Gaps 1;

Qy 1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTK-DQGTIDKIFQN 59
Db 7 SQLERSISTIINVFHQYSRKYGHPDTLNKAEFKEMVNKDLPNFLKREKRNENLLRDIMED 66

Qy 60 LDANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
Db 67 LDTNQDNQLSFEECMMLMGKLIFACHEKLHE 97

RESULT 13
US-09-492-026-7
; Sequence 7, Application US/09492026A
; Publication No. US20030096337A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; Bandman, Olga
; Corley, Neil C.
; Lal, Preeti
; Shah, Purvi
; TITLE OF INVENTION: HUMAN S100 PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/492,026A
; FILING DATE: 26-Jan-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Colette C. Muenzen
; REGISTRATION NUMBER: 39,784
; REFERENCE/DOCKET NUMBER: PF-0373 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 113 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 488157
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-492-026-7

Query Match 43.8%; Score 204.5; DB 10; Length 113;
Best Local Similarity 41.8%; Pred. No. 7.4e-15;

```
 Matches   38;  Conservative   24;  Mismatches   28;  Indels    1;  Gaps    1;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTK-DQGTIDKIFQN 59
              ::||  :  |||:||||| : || ||| | | |::: |:||| ||  | ::  :  | ::
Db          7 SQLERSISTIINVFHQYSRKYGHPDTLNKAEFKEMVNKDLPNFLKREKRNENLLRDIMED 66

Qy         60 LDANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
              || ||| |:||:| ::|:  ::   |: :|:
Db         67 LDTNQDNQLSFEECMMLMGKLIFACHEKLHE 97


RESULT 14
US-09-864-761-41579
  Sequence 41579, Application US/09864761
  Patent No. US20020048763A1
  GENERAL INFORMATION:
   APPLICANT: Penn, Sharron G.
   APPLICANT:  Rank, David R.
   APPLICANT:  Hanzel, David K.
   APPLICANT:  Chen, Wensheng
   TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
   TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
   FILE REFERENCE: Aeomica-X-1
   CURRENT APPLICATION NUMBER: US/09/864,761
   CURRENT FILING DATE:  2001-05-23
   PRIOR APPLICATION NUMBER: US 60/180,312
   PRIOR FILING DATE: 2000-02-04
   PRIOR APPLICATION NUMBER: US 60/207,456
   PRIOR FILING DATE: 2000-05-26
   PRIOR APPLICATION NUMBER: US 09/632,366
   PRIOR FILING DATE: 2000-08-03
   PRIOR APPLICATION NUMBER: GB 24263.6
   PRIOR FILING DATE: 2000-10-04
   PRIOR APPLICATION NUMBER: US 60/236,359
   PRIOR FILING DATE: 2000-09-27
   PRIOR APPLICATION NUMBER: PCT/US01/00666
   PRIOR FILING DATE: 2001-01-30
   PRIOR APPLICATION NUMBER: PCT/US01/00667
   PRIOR FILING DATE: 2001-01-30
   PRIOR APPLICATION NUMBER: PCT/US01/00664
   PRIOR FILING DATE: 2001-01-30
   PRIOR APPLICATION NUMBER: PCT/US01/00669
   PRIOR FILING DATE: 2001-01-30
   PRIOR APPLICATION NUMBER: PCT/US01/00665
   PRIOR FILING DATE: 2001-01-30
   PRIOR APPLICATION NUMBER: PCT/US01/00668
   PRIOR FILING DATE: 2001-01-30
   PRIOR APPLICATION NUMBER: PCT/US01/00663
   PRIOR FILING DATE: 2001-01-30
   PRIOR APPLICATION NUMBER: PCT/US01/00662
   PRIOR FILING DATE: 2001-01-30
   PRIOR APPLICATION NUMBER: PCT/US01/00661
   PRIOR FILING DATE: 2001-01-30
   PRIOR APPLICATION NUMBER: PCT/US01/00670
   PRIOR FILING DATE: 2001-01-30
   PRIOR APPLICATION NUMBER: US 60/234,687
   PRIOR FILING DATE: 2000-09-21
   PRIOR APPLICATION NUMBER: US 09/608,408
   PRIOR FILING DATE: 2000-06-30
   PRIOR APPLICATION NUMBER: US 09/774,203
   PRIOR FILING DATE: 2001-01-29
   NUMBER OF SEQ ID NOS: 49117
   SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
  SEQ ID NO 41579
    LENGTH: 46
    TYPE: PRT
    ORGANISM: Homo sapiens
    FEATURE:
    OTHER INFORMATION: MAP TO AC011666.18
    OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
    OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
    OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
    OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
    OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
    OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
    OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 7.1
    OTHER INFORMATION: EST_HUMAN HIT: AV715719.1, EVALUE 1.00e-19
    OTHER INFORMATION: SWISSPROT HIT: P80511, EVALUE 1.00e-20
  US-09-864-761-41579

    Query Match             41.1%;  Score 192;  DB 9;  Length 46;
    Best Local Similarity   80.0%;  Pred. No. 5.7e-14;
    Matches   36;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

  Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLK 45
                |||||:|||||:|||||||||| ||:||| | |||||:|||| ||:|
  Db          2 TKLEEHLEGIVNIFHQYSVRKGHFDTLSKGELKQLLTKELANTIK 46


  RESULT 15
  US-09-919-172-102
  ; Sequence 102, Application US/09919172
  ; Patent No. US20020119463A1
  ; GENERAL INFORMATION:
  ;  APPLICANT: Faris, Mary
  ;  APPLICANT:  Turner, Christopher M.
  ;  TITLE OF INVENTION: PROSTATE CANCER MARKERS
  ;  FILE REFERENCE: PA-0036 US
  ;  CURRENT APPLICATION NUMBER: US/09/919,172
  ;  CURRENT FILING DATE:  2001-07-30
  ;  PRIOR APPLICATION NUMBER: 60/222,469
  ;  PRIOR FILING DATE: 2000-07-28
  ;  NUMBER OF SEQ ID NOS: 102
  ;  SOFTWARE: PERL Program
  ; SEQ ID NO 102
  ;   LENGTH: 95
  ;   TYPE: PRT
  ;   ORGANISM: Homo sapiens
  ;   FEATURE:
  ;   NAME/KEY: misc_feature
  ;   OTHER INFORMATION: Incyte ID No. US20020119463A1 1422432CD1
  US-09-919-172-102

    Query Match             38.8%;  Score 181;  DB 9;  Length 95;
    Best Local Similarity   43.3%;  Pred. No. 2.2e-12;
    Matches   39;  Conservative   15;  Mismatches   36;  Indels    0;  Gaps    0;

  Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
                |:||  :  ||::| :||   |   || | ||| |: ||||  |:: ||:  :||: ::|
  Db          2 TELETAMGMIIDVFSRYSGSEGSTQTLTKGELKVLMEKELPGFLQSGKDKDAVDKLLKDL 61

  Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
                ||| | || | ||:| |  :   |   |
  Db         62 DANGDAQVDFSEFIVFVAAITSACHKYFEK 91


  Search completed: March  5, 2004, 09:51:39
  Job time : 35 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

)M protein - protein search, using sw model

Run on: March 5, 2004, 09:43:38 ; Search time 20 Seconds
(without alignments)
437.671 Million cell updates/sec

Title: US-09-646-651C-1
Perfect score: 467
Sequence: 1 TKLEDHLEGIINIFHQYSVR.....EFVVLVTDVLITAHDNIHKE 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	467	100.0	91	2	A55406	calgranulin c - pi
2	332	71.1	92	2	JC4712	S-100 calcium-bind
3	261.5	56.0	122	1	A42628	calgranulin B - bo
4	213.5	45.7	114	1	B31848	calgranulin B [val
5	204.5	43.8	113	1	JN0686	calgranulin B - ra
6	181	38.8	95	2	S24146	S-100 protein P -
7	178.5	38.2	113	1	S68242	calgranulin B - mo
8	167	35.8	95	1	S35985	S-100 protein alph
9	164	35.1	92	2	A26557	S-100 protein beta
10	163	34.9	89	1	I56163	calgranulin A - mo
11	163	34.9	91	1	BCBOIB	S-100 protein beta
12	163	34.9	92	2	A48015	S-100 protein beta
13	163	34.9	101	2	S06207	calvasculin - mous
14	162	34.7	92	1	BCHUIB	S-100 protein beta
15	161	34.5	94	1	BCBOIA	S-100 protein alph
16	160	34.3	94	1	BCHUIA	S-100 protein alph
17	160	34.3	101	2	S01759	calvasculin - rat
18	159	34.0	89	1	JN0685	calgranulin A - ra
19	156	33.4	101	2	A48219	calvasculin - huma
20	154	33.0	102	1	JQ1300	calgizzarin - rabb
21	153	32.8	100	2	A53217	placental calcium-
22	151	32.3	93	1	BCHUCF	calgranulin A [val
23	151	32.3	591	2	A45135	profilaggrin - hum
24	150	32.1	105	1	I37080	calgizzarin - huma
25	148	31.7	98	2	A41988	S-100 calcium-bind
26	147	31.5	306	2	A48118	major epidermal ca
27	144.5	30.9	79	1	KLPGI	calcium-binding pr
28	143.5	30.7	110	2	B48219	S-100 calcium-bind
29	140.5	30.1	98	2	JC5064	S-100 calcium-bind
30	139.5	29.9	79	1	JN0246	calcium-binding pr
31	138	29.6	97	2	A30129	S-100 protein, lun
32	136	29.1	90	1	BCHUY	calcyclin - human
33	135.5	29.0	97	2	A28489	calpactin I light
34	135	28.9	79	1	KLBOI	calcium-binding pr
35	135	28.9	90	1	S27011	calcyclin - rabbit
36	134.5	28.8	95	2	A31373	calpactin I light
37	133.5	28.6	97	1	JH0663	calpactin I light
38	133	28.5	99	2	S20342	calcium-binding pr
39	131	28.1	89	2	A54314	calcyclin - mouse
40	129.5	27.7	95	1	LUPG10	calpactin I light
41	129.5	27.7	97	2	JC1139	calpactin I light
42	129.5	27.7	97	2	B28489	calpactin I light
43	128	27.4	90	2	B28363	calcyclin - rat
44	127.5	27.3	217	2	JE0330	26-kDa Ca2+-bindin
45	125.5	26.9	98	2	JC5065	calcium-binding pr

ALIGNMENTS

RESULT 1
A55406
calgranulin c - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 23-May-1997
C;Accession: A55406
R;Dell'Angelica, E.C.; Schleicher, C.H.; Santome, J.A.
J. Biol. Chem. 269, 28929-28936, 1994
A;Title: Primary structure and binding properties of calgranulin C, a novel S100-like ca
A;Reference number: A55406; MUID:95050708; PMID:7961855
A;Accession: A55406
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-91 <DEL>
C;Superfamily: S-100 protein; calmodulin repeat homology
C;Keywords: calcium binding; EF hand
F;48-80/Domain: calmodulin repeat homology <EF2>

Query Match 100.0%; Score 467; DB 2; Length 91;
Best Local Similarity 100.0%; Pred. No. 1.6e-37;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              |||||||||||||||||||||||||||||||
Db         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
```

RESULT 2
JC4712
S-100 calcium-binding protein A12 - human
N;Alternate names: calcium-binding amniotic fluid protein 1 (CAAF1); calgranulin C; calg
utrophil protein
C;Species: Homo sapiens (man)
C;Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 13-Aug-1999
C;Accession: JC4712; JC4717; JC4891; S56113; S56114
R;Yamamura, T.; Hitomi, J.; Nagasaki, K.; Suzuki, M.; Takahashi, E.; Saito, S.; Tsukada,
Biochem. Biophys. Res. Commun. 221, 356-360, 1996
A;Title: Human CAAF1 gene - molecular cloning, gene structure, and chromosome mapping.
A;Reference number: JC4712; MUID:96192053; PMID:8619860
A;Accession: JC4712
A;Molecule type: mRNA
A;Residues: 1-92 <YAM>
A;Cross-references: DDBJ:D83657; NID:g1502284; PIDN:BAA12030.1; PID:g1502285
R;Marti, T.; Erttmann, K.D.; Gallin, M.Y.
Biochem. Biophys. Res. Commun. 221, 454-458, 1996
A;Title: Host-parasite interaction in human onchocerciasis: Identification and sequence
A;Reference number: JC4717; MUID:96192069; PMID:8619876
A;Accession: JC4717

A;Molecule type: protein
A;Residues: 2-92 <MAR>
A;Experimental source: Onchocerca volvulus infecting human tissue
R;Ilg, E.C.; Troxler, H.; Buergisser, D.M.; Kuster, T.; Markert, M.; Guignard, F.; Hunzi
Biochem. Biophys. Res. Commun. 225, 146-150, 1996
A;Title: Amino acid sequence determination of human S100 A12 (P6, Calgranulin C, CGRP, C
A;Reference number: JC4891; MUID:96332419; PMID:8769108
A;Accession: JC4891
A;Molecule type: protein
A;Residues: 2-92 <ILG>
R;Guignard, F.; Mauel, J.; Markert, M.
Biochem. J. 309, 395-401, 1995
A;Title: Identification and characterization of a novel human neutrophil protein related
A;Reference number: S56113; MUID:95351965; PMID:7626002
A;Accession: S56113
A;Status: preliminary
A;Molecule type: protein
A;Residues: 'XX',4-14,'X',16-17,'XXXX' <GUI1>
A;Experimental source: isoform 6a
A;Accession: S56114
A;Status: preliminary
A;Molecule type: protein
A;Residues: 2-21 <GUI2>
A;Experimental source: isoform 6b
C;Comment: This protein is released by activated neutrophils in the course of inflammato
C;Genetics:
A;Gene: GDB:S100A12; p6; MRP6; CGRP; CAAF1
A;Cross-references: GDB:5218374
A;Map position: 1q21-1q21
C;Complex: monomer
C;Superfamily: S-100 protein; calmodulin repeat homology
C;Keywords: calcium binding; EF hand; monomer; neutrophil; zinc
F;2-92/Product: S-100 calcium-binding protein A12 #status experimental <MAT>
F;6-39/Domain: calmodulin repeat homology <EF1>
F;49-81/Domain: calmodulin repeat homology <EF2>
F;86-90/Region: zinc binding #status predicted

  Query Match 71.1%; Score 332; DB 2; Length 92;
  Best Local Similarity 70.3%; Pred. No. 1e-24;
  Matches 64; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

```
Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              |||:|||||:|||||||||| ||:||| | |||||:|||| ||:|| ||:  ||:||| |
Db          2 TKLEEHLEGIVNIFHQYSVRKGHFDTLSKGELKQLLTKELANTIKNIKDKAVIDEIFQGL 61

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              |||||||| |:||: ||   |  || : |||
Db         62 DANQDEQVDFQEFISLVAIALKAAHYHTHKE 92
```

RESULT 3
A42628
calgranulin B - bovine (fragment)
N;Alternate names: calcium-binding protein MRP-14; macrophage migration inhibitory facto
in 2
C;Species: Bos primigenius taurus (cattle)
C;Date: 30-Sep-1993 #sequence_revision 23-May-1997 #text_change 23-May-1997
C;Accession: B22309; A42628
R;Tang, T.K.; Hong, T.M.; Lin, C.Y.; Lai, M.L.; Liu, C.H.L.; Lo, H.J.; Wang, M.E.; Chen,
submitted to the Protein Sequence Database, July 1992
A;Reference number: A22309
A;Accession: B22309
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-122 <TAN>
R;Dianoux, A.C.; Stasia, M.J.; Garin, J.; Gagnon, J.; Vignais, P.V.
Biochemistry 31, 5898-5905, 1992
A;Title: The 23-kilodalton protein, a substrate of protein kinase C, in bovine neutrophi
A;Reference number: A42628; MUID:92304974; PMID:1610833
A;Accession: A42628
A;Molecule type: protein
A;Residues: 4-32,'F',34-56 <DIA>
C;Complex: heterodimer and higher complexes with calgranulin A
C;Superfamily: S-100 protein; calmodulin repeat homology
C;Keywords: blocked amino end; calcium binding; EF hand; heterodimer; inflammation; pho
F;6-40/Domain: calmodulin repeat homology <EF1>
F;50-82/Domain: calmodulin repeat homology <EF2>

  Query Match 56.0%; Score 261.5; DB 1; Length 122;
  Best Local Similarity 53.3%; Pred. No. 6.6e-18;
  Matches 48; Conservative 23; Mismatches 18; Indels 1; Gaps 1;

```
Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTL-KNTKDQGTIDKIFQN 59
              :::|  :| |||||||||||||||||::| |||: |||| | |  |::  |::| ::
Db          2 SQMESSIETIINIFHQYSVRLGHYDTLIQKESKQLVQKELPNFLKKQKKNEAAINEIMED 61

Qy         60 LDANQDEQVSFKEFVVLVTDVLITAHDNIH 89
              || | |:|:||:||::||  : : :|: :|
Db         62 LDTNVDKQLSFEEFIMLVARLTVASHEEMH 91
```

RESULT 4
B31848
calgranulin B [validated] - human
N;Alternate names: calcium-binding protein MRP-14; cystic fibrosis-associated antigen (
(MRP-14); MIF-related 14K protein; S-100 calcium-binding protein A9 (S100A9)
C;Species: Homo sapiens (man)
C;Date: 21-May-1990 #sequence_revision 23-May-1997 #text_change 08-Dec-2000
C;Accession: B31848; S00667; A33819; B60911; B61082; D54327
R;Lagasse, E.; Clerc, R.G.
Mol. Cell. Biol. 8, 2402-2410, 1988
A;Title: Cloning and expression of two human genes encoding calcium-binding proteins th
A;Reference number: A93102; MUID:88302148; PMID:3405210
A;Accession: B31848
A;Molecule type: DNA
A;Residues: 1-114 <LAG>
A;Cross-references: GB:M21064; NID:g188689; PIDN:AAA36326.1; PID:g386958
R;Odink, K.; Cerletti, N.; Brueggen, J.; Clerc, R.G.; Tarcsay, L.; Zwadlo, G.; Gerhards
Nature 330, 80-82, 1987
A;Title: Two calcium-binding proteins in infiltrate macrophages of rheumatoid arthritis
A;Reference number: S00667; MUID:88039099; PMID:3313057
A;Accession: S00667
A;Molecule type: mRNA
A;Residues: 1-114 <ODI>
A;Cross-references: EMBL:X06233; NID:g34770; PIDN:CAA29579.1; PID:g34771
A;Note: parts of this sequence were confirmed by protein sequencing
R;Murao, S.; Collart, F.R.; Huberman, E.
J. Biol. Chem. 264, 8356-8360, 1989
A;Title: A protein containing the cystic fibrosis antigen is an inhibitor of protein ki
A;Reference number: A33819; MUID:89255276; PMID:2656677
A;Accession: A33819
A;Molecule type: mRNA
A;Residues: 1-114 <MUR>
A;Cross-references: GB:M26311; NID:g862619; PIDN:AAA68480.1; PID:g516621
A;Note: part of this sequence was confirmed by protein sequencing; the amino end of the
R;Andersson, K.B.; Sletten, K.; Berntzen, H.B.; Dale, I.; Brandtzaeg, P.; Jellum, E.; F
Scand. J. Immunol. 28, 241-245, 1988
A;Title: The leucocyte L1 protein: identity with the cystic fibrosis antigen and the ca
A;Reference number: A60911; MUID:88321575; PMID:3413449
A;Accession: B60911
A;Molecule type: protein
A;Residues: 39-42,'X',44-50;64-77,'X',79;84,'X',86-90,'X',92-94,'X',96-98 <AND>
R;Tobe, T.; Murakami, K.; Tomita, M.; Nozawa, R.
Chem. Pharm. Bull. 37, 1576-1580, 1989
A;Title: Amino acid sequences of 60B8 antigens induced in HL-60 cells by 1,25-dihydroxy
A;Reference number: A61082; MUID:89376638; PMID:2776242
A;Accession: B61082
A;Molecule type: protein
A;Residues: 5-77;80-90,'A',92-114 <TOB>
A;Note: the blocked amino end of the mature protein is identified as 2-Thr; residue 91-1
R;Madsen, P.; Rasmussen, H.H.; Leffers, H.; Honore, B.; Dejgaard, K.; Olsen, E.; Kiil,
E.
J. Invest. Dermatol. 97, 701-712, 1991
A;Title: Molecular cloning, occurrence, and expression of a novel partially secreted pr
A;Reference number: A54327; MUID:92043866; PMID:1940442
A;Accession: D54327

```
;Molecule type: protein
;Residues: 11-19;26-38;94-105,'X',107 <MAD>
;Note: in several peptide samples no PTH was detected for 95-His but in one peptide PTH
;Comment: This protein appears to be expressed only in cells of myeloid origin actively
;Comment: The presence of 3'-methylhistidine at position 105, corresponding to 107-His
;Genetics:
;Gene: GDB:S100A9; 60B8AG; CAGB; CFAG; LIAG; MAC387; MIF; MRP14; NIF; P14
;Cross-references: GDB:120570; OMIM:123886
;Map position: 1q21-1q21
;Introns: 50/3
;Note: the first intron occurs before the initiator codon
;Complex: heterodimer and higher complexes with calgranulin A (see PIR:BCHUCF)
;Superfamily: S-100 protein; calmodulin repeat homology
;Keywords: blocked amino end; calcium binding; EF hand; heterodimer; inflammation; phos
;2-114/Product: calgranulin B #status experimental <MAT>
;10-44/Domain: calmodulin repeat homology <EF1>
;54-86/Domain: calmodulin repeat homology <EF2>
;2/Modified site: blocked amino end (Thr) (in mature form) (probably acetylated) #statu
;113/Binding site: phosphate (Thr) (covalent) #status predicted

  Query Match             45.7%;  Score 213.5;  DB 1;  Length 114;
  Best Local Similarity   45.1%;  Pred. No. 2.1e-13;
  Matches   41;  Conservative   25;  Mismatches   24;  Indels    1;  Gaps    1;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTL-KNTKDQGTIDKIFQN 59
              ::|| ::| ||| |||||||||||| ||| : | |:|: |:| | | |  |:: |: | ::
Db          6 SQLERNIETIINTFHQYSVKLGHPDTLNQGEFKELVRKDLQNFLKKENKNEKVIEHIMED 65

Qy         60 LDANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
              || | |:|:||:||::|:  :   :|: :|:
Db         66 LDTNADKQLSFEEFIMLMARLTWASHEKMHE 96


RESULT 5
JN0686
algranulin B - rat
;Alternate names: calcium-binding protein MRP-14; macrophage migration inhibitory facto
;Species: Rattus norvegicus (Norway rat)
;Date: 03-May-1994 #sequence_revision 23-May-1997 #text_change 22-Jun-1999
;Accession: JN0686
;Imamichi, T.; Uchida, I.; Wahl, S.M.; McCartney-Francis, N.
Biochem. Biophys. Res. Commun. 194, 819-825, 1993
;Title: Expression and cloning of migration inhibitory factor-related protein (MRP)8 an
;Reference number: JN0685; MUID:93343942; PMID:8343166
;Accession: JN0686
;Molecule type: mRNA
;Residues: 1-113 <IMA>
;Cross-references: GB:L18948; NID:g488156; PIDN:AAA18214.1; PID:g488157
;Genetics:
;Gene: MRP14
;Complex: heterodimer and higher complexes with calgranulin A
;Superfamily: S-100 protein; calmodulin repeat homology
;Keywords: acetylated amino end; calcium binding; EF hand; heterodimer; inflammation; p
;2-113/Product: calgranulin B #status predicted <MAT>
;11-45/Domain: calmodulin repeat homology <EF1>
;55-87/Domain: calmodulin repeat homology <EF2>
;2/Modified site: acetylated amino end (Ala) (in mature form) #status predicted
;80-91/Disulfide bonds: #status predicted

  Query Match             43.8%;  Score 204.5;  DB 1;  Length 113;
  Best Local Similarity   41.8%;  Pred. No. 1.5e-12;
  Matches   38;  Conservative   24;  Mismatches   28;  Indels    1;  Gaps    1;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTK-DQGTIDKIFQN 59
              ::||  :  |||:||||| : || ||| | | |::: |:||| ||  | ::  :  | ::
Db          7 SQLERSISTIINVFHQYSRKYGHPDTLNKAEFKEMVNKDLPNFLKREKRNENLLRDIMED 66

Qy         60 LDANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
              || ||| |:||:| ::|:  ::   |: :|:
Db         67 LDTNQDNQLSFEECMMLMGKLIFACHEKLHE 97


RESULT 6
S24146
S-100 protein P - human
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 13-Aug-1999
C;Accession: S24146; PS0340
R;Becker, T.; Gerke, V.; Kube, E.; Weber, K.
Eur. J. Biochem. 207, 541-547, 1992
A;Title: S100P, a novel Ca(2+)-binding protein from human placenta. cDNA cloning, recomb
A;Reference number: S24146; MUID:92339442; PMID:1633809
A;Accession: S24146
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-95 <BEC>
A;Cross-references: EMBL:X65614; NID:g36177; PIDN:CAA46566.1; PID:g36178
R;Emoto, Y.; Kobayashi, R.; Akatsuka, H.; Hidaka, H.
Biochem. Biophys. Res. Commun. 182, 1246-1253, 1992
A;Title: Purification and characterization of a new member of the S-100 protein family f
A;Reference number: PS0340; MUID:92171935; PMID:1540168
A;Accession: PS0340
A;Molecule type: protein
A;Residues: 1-31,'T',33-84,'X',86-91 <EMO>
A;Experimental source: placenta
C;Genetics:
A;Gene: GDB:S100P
A;Cross-references: GDB:134405; OMIM:600614
A;Map position: 4p16-4p16
C;Superfamily: S-100 protein; calmodulin repeat homology
C;Keywords: calcium binding; EF hand; placenta
F;6-40/Domain: calmodulin repeat homology <EF1>
F;49-81/Domain: calmodulin repeat homology <EF2>

  Query Match             38.8%;  Score 181;  DB 2;  Length 95;
  Best Local Similarity   43.3%;  Pred. No. 2.1e-10;
  Matches   39;  Conservative   15;  Mismatches   36;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              |:||  :  ||::| :||   |  || | ||| |: |||| |:: ||:  :||: ::|
Db          2 TELETAMGMIIDVFSRYSGSEGSTQTLTKGELKVLMEKELPGFLQSGKDKDAVDKLLKDL 61

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
              ||| | || | ||:| |  :    |    |
Db         62 DANGDAQVDFSEFIVFVAAITSACHKYFEK 91


RESULT 7
S68242
calgranulin B - mouse
N;Alternate names: calcium-binding protein MRP-14; macrophage migration inhibitory facto
C;Species: Mus musculus (house mouse)
C;Date: 06-Sep-1996 #sequence_revision 23-May-1997 #text_change 22-Jun-1999
C;Accession: S68242; S68272
R;Lagasse, E.; Weissman, I.L.
submitted to the EMBL Data Library, February 1992
A;Description: Mouse MRP8 and MRP14, two intracellular calcium-binding proteins associat
A;Reference number: S68242
A;Accession: S68242
A;Molecule type: mRNA
A;Residues: 1-113 <LAG>
A;Cross-references: EMBL:M83219; NID:g199807; PIDN:AAB07228.1; PID:g199808
R;Raftery, M.J.; Harrison, C.A.; Alewood, P.; Jones, A.; Geczy, C.L.
Biochem. J. 316, 285-293, 1996
A;Title: Isolation of the murine S100 protein MRP14 (14 kDa migration-inhibitory-factor-
ding.
A;Reference number: S68272; MUID:96235204; PMID:8645219
A;Accession: S68272
A;Molecule type: protein
A;Residues: 2-10;95-109 <RAF>
A;Note: 107-His is identified as 3'-methylhistidine; the authors' source for the referen
ylhistidine
C;Genetics:
A;Gene: MRP14
C;Complex: heterodimer and higher complexes with calgranulin A
C;Superfamily: S-100 protein; calmodulin repeat homology
```

```
C;Keywords: acetylated amino end; calcium binding; EF hand; heterodimer; inflammation; m
F;2-113/Product: calgranulin B #status predicted <MAT>
F;11-45/Domain: calmodulin repeat homology <EF1>
F;55-87/Domain: calmodulin repeat homology <EF2>
F;2/Modified site: acetylated amino end (Ala) (in mature form) #status experimental
F;80-91/Disulfide bonds: #status experimental
F;103,105,107/Binding site: zinc (His) #status predicted
F;107/Modified site: 3'-methylhistidine (His) #status experimental

  Query Match             38.2%;  Score 178.5;  DB 1;  Length 113;
  Best Local Similarity   36.3%;  Pred. No. 4.4e-10;
  Matches   33;  Conservative   27;  Mismatches   30;  Indels    1;  Gaps    1;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTK-DQGTIDKIFQN 59
              :::|  :  ||: |||||  : || ||| |:| :|::  :|   :|  | ::  |: | ::
Db          7 SQMERSITTIIDTFHQYSRKEGHPDTLSKKEFRQMVEAQLATFMKKEKRNEALINDIMED 66

Qy         60 LDANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
              || ||| |:||:| ::|:  ::    |: :|:
Db         67 LDTNQDNQLSFEECMMLMAKLIFACHEKLHE 97


RESULT 8
S35985
S-100 protein alpha chain - weatherfish
C;Species: Misgurnus fossilis (weatherfish)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: S35985
R;Ivanenkov, V.V.; Gerke, V.; Minin, A.A.; Plessmann, U.; Weber, K.
Mech. Dev. 42, 151-158, 1993
A;Title: Transduction of Ca(2+) signals upon fertilization of eggs; identification of an
A;Reference number: S35985; MUID:94031845; PMID:8217841
A;Accession: S35985
A;Molecule type: protein
A;Residues: 1-95 <IVA>
C;Superfamily: S-100 protein; calmodulin repeat homology
C;Keywords: calcium binding; EF hand
F;49-81/Domain: calmodulin repeat homology <EF2>

  Query Match             35.8%;  Score 167;  DB 1;  Length 95;
  Best Local Similarity   40.7%;  Pred. No. 4.4e-09;
  Matches   35;  Conservative   18;  Mismatches   33;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ::||  :| :| :|| || : |    | | ||| |:  || : |  :||    ::||  :|
Db          2 SQLESAMESLIKVFHTYSSKEGDKYKLSKAELKSLLQGELNDFLSASKDPMVVEKIMSDL 61

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHD 86
              | ||| :| |:||||||  : :  ::
Db         62 DENQDGEVDFQEFVVLVAALTVACNE 87


RESULT 9
A26557
S-100 protein beta chain - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 25-Mar-1988 #sequence_revision 04-Nov-1994 #text_change 13-Aug-1999
C;Accession: A60046; S07357; A26557
R;Maeda, T.; Usui, H.; Araki, K.; Kuwano, R.; Takahashi, Y.; Suzuki, Y.
Brain Res. Mol. Brain Res. 10, 193-202, 1991
A;Title: Structure and expression of rat S-100 beta subunit gene.
A;Reference number: A60046; MUID:91359841; PMID:1653388
A;Accession: A60046
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-92 <MAE>
A;Cross-references: GB:S53527
R;Kuwano, R.; Usui, H.; Maeda, T.; Fukui, T.; Yamanari, N.; Ohtsuka, E.; Ikehara, M.; Ta
Nucleic Acids Res. 12, 7455-7465, 1984
A;Title: Molecular cloning and the complete nucleotide sequence of cDNA to mRNA for S-10
A;Reference number: S07357; MUID:85037924; PMID:6093041
A;Accession: S07357
A;Molecule type: mRNA
A;Residues: 1-92 <KUW>
A;Cross-references: EMBL:X01090; NID:g57174; PIDN:CAA25567.1; PID:g57175
R;Dunn, R.; Landry, C.; O'Hanlon, D.; Dunn, J.; Allore, R.; Brown, I.; Marks, A.
J. Biol. Chem. 262, 3562-3566, 1987
A;Title: Reduction in S100 protein Beta-subunit mRNA in C6 rat glioma cells following t
A;Reference number: A26557; MUID:87137648; PMID:3818655
A;Accession: A26557
A;Molecule type: mRNA
A;Residues: 6-92 <DUN>
A;Cross-references: GB:M15705
C;Comment: S-100 protein occurs as alpha-beta heterodimers, alpha-alpha heterodimers, a
C;Genetics:
A;Introns: 46/3
C;Superfamily: S-100 protein; calmodulin repeat homology
C;Keywords: brain; calcium binding; dimer; EF hand; zinc
F;2-92/Product: S-100 protein beta chain #status predicted <MAT>
F;6-40/Domain: calmodulin repeat homology <EF1>

  Query Match             35.1%;  Score 164;  DB 2;  Length 92;
  Best Local Similarity   37.4%;  Pred. No. 8.2e-09;
  Matches   34;  Conservative   19;  Mismatches   38;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ::||  :  :|::|||||  | |    | | |||:||  || : |:  |:|  :||: : |
Db          2 SELEKAMVALIDVFHQYSGREGDKHKLKKSELKELINNELSHFLEEIKEQEVVDKVMETL 61

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              | : | :  |:||:  |: |    |:    |
Db         62 DEDGDGECDFQEFMAFVSMVTTACHEFFEHE 92


RESULT 10
I56163
calgranulin A - mouse
N;Alternate names: calcium-binding protein MRP-8; CP-10 chemotactic protein; macrophage
C;Species: Mus musculus (house mouse)
C;Date: 26-Jul-1996 #sequence_revision 23-May-1997 #text_change 22-Jun-1999
C;Accession: I56163; A42488
R;Lackmann, M.; Rajasekariah, P.; Iismaa, S.E.; Jones, G.; Cornish, C.J.; Hu, S.; Simps
J. Immunol. 150, 2981-2991, 1993
A;Title: Identification of a chemotactic domain of the pro-inflammatory S100 protein CP
A;Reference number: I56163; MUID:93203618; PMID:8454868
A;Accession: I56163
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-89 <LAC1>
A;Cross-references: GB:S57123; NID:g298706; PIDN:AAB25840.1; PID:g298707
R;Lackmann, M.; Cornish, C.J.; Simpson, R.J.; Moritz, R.L.; Geczy, C.L.
J. Biol. Chem. 267, 7499-7504, 1992
A;Title: Purification and structural analysis of a murine chemotactic cytokine (CP-10)
A;Reference number: A42488; MUID:92218405; PMID:1559987
A;Accession: A42488
A;Molecule type: protein
A;Residues: 2-77 <LAC2>
A;Note: sequence extracted from NCBI backbone (NCBIP:94068)
C;Genetics:
A;Gene: MRP8
C;Complex: homodimer; heterodimer and higher complexes with calgranulin B
C;Superfamily: S-100 protein; calmodulin repeat homology
C;Keywords: calcium binding; cytokine; EF hand; heterodimer; homodimer; inflammation
F;2-89/Product: calgranulin A #status experimental <MAT>
F;7-41/Domain: calmodulin repeat homology <EF1>
F;46-78/Domain: calmodulin repeat homology <EF2>

  Query Match             34.9%;  Score 163;  DB 1;  Length 89;
  Best Local Similarity   33.0%;  Pred. No. 9.8e-09;
  Matches   30;  Conservative   25;  Mismatches   32;  Indels    4;  Gaps    1;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ::||  |  :|:::| ||   |::  | | : |:::| | :|  ::|      |: :|: |
Db          3 SELEKALSNLIDVYHNYSNIQGNHHALYKNDFKKMVTTECPQFVQNI----NIENLFREL 58

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
```

```
                  | | |  ::|:||: :|  | : :| : |||
Db          59 DINSDNAINFEEFLAMVIKVGVASHKDSHKE 89
```

RESULT 11
3CBOIB
S-100 protein beta chain - bovine
N;Alternate names: neurocalcin delta-binding protein S100-beta
C;Species: Bos primigenius taurus (cattle)
C;Date: 31-May-1979 #sequence_revision 14-Nov-1983 #text_change 24-Nov-1999
C;Accession: A91254; B91110; A90075; S54348; A03077
R;Isobe, T.; Okuyama, T.
Eur. J. Biochem. 89, 379-388, 1978
A;Title: The amino-acid sequence of S-100 protein (PAP-I-b protein) and its relation to
A;Reference number: A91254; MUID:79045265; PMID:710399
A;Accession: A91254
A;Molecule type: protein
A;Residues: 'ESEL',5-91 <ISO>
A;Experimental source: brain
A;Note: this sequence has since been revised in reference A91110
R;Isobe, T.; Okuyama, T.
Eur. J. Biochem. 116, 79-86, 1981
A;Title: The amino-acid sequence of the alpha subunit in bovine brain S-100 a protein.
A;Reference number: A91110; MUID:81236562; PMID:7250124
A;Accession: B91110
A;Molecule type: protein
A;Residues: 1-91 <IS2>
R;Baudier, J.; Gerard, D.
Biochemistry 22, 3360-3369, 1983
A;Title: Ions binding to S100 proteins: structural changes induced by calcium and zinc c
A;Reference number: A90471; MUID:84000339; PMID:6615778
A;Contents: annotation; metal ion-binding properties
R;Marshak, D.R.; Umekawa, H.; Watterson, D.M.; Hidaka, H.
Arch. Biochem. Biophys. 240, 777-780, 1985
A;Title: Structural characterization of the calcium binding protein S100 from adipose ti
A;Reference number: A90075; MUID:85278169; PMID:4026304
A;Accession: A90075
A;Molecule type: protein
A;Residues: 1-91 <MAR>
A;Experimental source: adipose tissue
R;Okazaki, K.; Obata, N.H.; Inoue, S.; Hidaka, H.
Biochem. J. 306, 551-555, 1995
A;Title: S100-beta is a target protein of neurocalcin delta, an abundant isoform in glia
A;Reference number: S54343; MUID:95194333; PMID:7887910
A;Accession: S54348
A;Molecule type: protein
A;Residues: 56-61,'N',63-79,'V' <OKA>
C;Comment: The S-100 protein is composed of two related polypeptide chains, alpha and be
brain proteins, S-100 is also found in a variety of other tissues.
C;Comment: S-100 is an intracellular protein that weakly binds calcium. It binds zinc ve
es with different affinities exist for both ions on each monomer. Physiological concentr
cium-binding sites.
C;Superfamily: S-100 protein; calmodulin repeat homology
C;Keywords: blocked amino end; brain; calcium binding; EF hand; zinc
F;5-39/Domain: calmodulin repeat homology <EF1>
F;48-80/Domain: calmodulin repeat homology <EF2>
F;1/Modified site: blocked amino end (Ser) (probably acetylated) #status experimental
F;18,21,23,26,31/Binding site: calcium (Ser, Glu, Asp, Lys, Glu) #status predicted
F;61,63,65,67,72/Binding site: calcium (Asp, Asp, Asp, Glu, Glu) #status predicted

  Query Match   34.9%; Score 163; DB 1; Length 91;
  Best Local Similarity   36.3%; Pred. No. 1e-08;
  Matches   33; Conservative   20; Mismatches   38; Indels   0; Gaps   0;

```
Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ::||  :  :|::|||||| | |    | | |||:||  || : |:  |:|  :||: : |
Db          1 SELEKAVVALIDVFHQYSGREGDKHKLKKSELKELINNELSHFLEEIKEQEVVDKVMETL 60

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              |:: | :  |:||:  |  :    |:    |
Db         61 DSDGDGECDFQEFMAFVAMITTACHEFFEHE 91
```

RESULT 12
A48015
S-100 protein beta chain - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jun-1995 #sequence_revision 02-Jun-1995 #text_change 24-Nov-1999
C;Accession: A48015
R;Jiang, H.; Shah, S.; Hilt, D.C.
J. Biol. Chem. 268, 20502-20511, 1993
A;Title: Organization, sequence, and expression of the murine S100beta gene. Transcripti
A;Reference number: A48015; MUID:93388628; PMID:8376406
A;Accession: A48015
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-92 <JIA>
A;Cross-references: GB:L22144; NID:g404768; PIDN:AAA03075.1; PID:g404769
C;Genetics:
A;Introns: 46/3
C;Superfamily: S-100 protein; calmodulin repeat homology
C;Keywords: acetylated amino end; calcium binding; EF hand
F;2-92/Product: S-100 protein beta chain #status predicted <MAT>
F;6-40/Domain: calmodulin repeat homology <EF1>
F;49-81/Domain: calmodulin repeat homology <EF2>
F;2/Modified site: acetylated amino end (Ser) (in mature form) #status predicted
F;20,22,24,27,32/Binding site: calcium (Gly, Glu, Asp, Lys, Glu) #status predicted
F;62,64,66,68,70,73/Binding site: calcium (Asp, Asp, Asp, Glu, Asp, Glu) #status predict

  Query Match   34.9%; Score 163; DB 2; Length 92;
  Best Local Similarity   37.4%; Pred. No. 1e-08;
  Matches   34; Conservative   18; Mismatches   39; Indels   0; Gaps   0;

```
Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ::||  :  :|::|||||| | |    | | |||:||  || : |:  |:|  :||: : |
Db          2 SELEKAMVALIDVFHQYSGREGDKHKLKKSELKELINNELSHFLEEIKEQEVVDKVMETL 61

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              | : | :  |:||:  |  |    |:    |
Db         62 DEDGDGECDFQEFMAFVAMVTTACHEFFEHE 92
```

RESULT 13
S06207
calvasculin - mouse
N;Alternate names: calcium-binding protein mts1; calcium-binding protein pEL98; placenta
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 20-Jun-2000
C;Accession: S06207; JH0097; S07981; A26803; A41411; I48674
R;Ebralidze, A.; Tulchinsky, E.; Grigorian, M.; Afanasyeva, A.; Senin, V.; Revazova, E.;
Genes Dev. 3, 1086-1093, 1989
A;Title: Isolation and characterization of a gene specifically expressed in different me

A;Reference number: S06207; MUID:89378739; PMID:2550322
A;Accession: S06207
A;Molecule type: mRNA
A;Residues: 1-101 <EBR>
A;Cross-references: EMBL:X16190; NID:g54926; PIDN:CAA34316.1; PID:g54927
R;Tulchinsky, E.M.; Grigorian, M.S.; Ebralidze, A.K.; Milshina, N.I.; Lukanidin, E.M.
Gene 87, 219-223, 1990
A;Title: Structure of gene mts1, transcribed in metastatic mouse tumor cells.
A;Reference number: JH0097; MUID:90236313; PMID:2332170
A;Accession: JH0097
A;Molecule type: DNA
A;Residues: 1-101 <TUL>
A;Cross-references: GB:M36578; GB:M36579
A;Experimental source: liver
R;Tulchinsky, B.
submitted to the EMBL Data Library, August 1989
A;Reference number: S07981
A;Accession: S07981
A;Molecule type: DNA
A;Residues: 1-47,'VSGSXFNG',56-57,'RTDEAA' <TU2>
A;Cross-references: EMBL:X16094; NID:g53249; PIDN:CAA34224.1; PID:g53250
R;Jackson-Grusby, L.L.; Swiergiel, J.; Linzer, D.I.H.
Nucleic Acids Res. 15, 6677-6690, 1987
```
```

A;Title: A growth-related mRNA in cultured mouse cells encodes a placental calcium bindi
A;Reference number: A26803; MUID:87316927; PMID:3628004
A;Accession: A26803
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-101 <JAC>
A;Cross-references: GB:X05835; NID:g50310; PIDN:CAA29282.1; PID:g50311
R;Goto, K.; Endo, H.; Fujiyoshi, T.
J. Biochem. 103, 48-53, 1988
A;Title: Cloning of the sequences expressed abundantly in established cell lines: identi
A;Reference number: A41411; MUID:88198109; PMID:3162911
A;Accession: A41411
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-101 <GOT>
A;Cross-references: GB:D00208; NID:g220569; PIDN:BAA00148.1; PID:g220570
R;Tulchinsky, E.; Kramerov, D.; Ford, H.L.; Reshetnyak, E.; Lukanidin, E.; Zain, S.
Oncogene 8, 79-86, 1993
A;Title: Characterization of a positive regulatory element in the mts1 gene.
A;Reference number: I48674; MUID:93141279; PMID:8423998
A;Accession: I48674
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-47,'VSGSXFNGQ',48-54 <RES>
A;Cross-references: EMBL:X16094; NID:g53249; PIDN:CAA34224.1; PID:g53250
C;Comment: Gene mts1 is expressed in metastatic cells.
C;Genetics:
A;Gene: mts1
A;Introns: 47/3
C;Superfamily: S-100 protein; calmodulin repeat homology
C;Keywords: calcium binding; cancer; EF hand
F;7-41/Domain: calmodulin repeat homology <EF1>
F;50-82/Domain: calmodulin repeat homology <EF2>

Query Match 34.9%; Score 163; DB 2; Length 101;
Best Local Similarity 36.9%; Pred. No. 1.1e-08;
Matches 31; Conservative 23; Mismatches 30; Indels 0; Gaps 0;

```
Qy          3 LEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNLDA 62
              ||: |: |:: ||:|| : |    | | |||:|:|:|||: |    |:    |:  |||:
Db          5 LEEALDVIVSTFHKYSGKEGDKFKLNKTELKELLTRELPSFLGKRTDEAAFQKVMSNLDS 64

Qy         63 NQDEQVSFKEFVVLVTDVLITAHD 86
              |:| :| |:|: | :: : :  ::
Db         65 NRDNEVDFQEYCVFLSCIAMMCNE 88
```

RESULT 14
BCHUIB
S-100 protein beta chain [validated] - human
N;Alternate names: neural S-100 calcium-binding protein beta
C;Species: Homo sapiens (man)
C;Date: 04-Dec-1986 #sequence_revision 06-Jan-1995 #text_change 08-Dec-2000
C;Accession: A38364; A92972; A03076
R;Allore, R.J.; Friend, W.C.; O'Hanlon, D.; Neilson, K.M.; Baumal, R.; Dunn, R.J.; Marks
J. Biol. Chem. 265, 15537-15543, 1990
A;Title: Cloning and expression of the human S100beta gene.
A;Reference number: A38364; MUID:90368757; PMID:2394738
A;Accession: A38364
A;Molecule type: DNA
A;Residues: 1-92 <ALL>
A;Cross-references: GB:J05600; GB:M59486; NID:g337726; GB:M59487; NID:g337727; GB:M59488
R;Jensen, R.; Marshak, D.R.; Anderson, C.; Lukas, T.J.; Watterson, D.M.
J. Neurochem. 45, 700-705, 1985
A;Title: Characterization of human brain S100 protein fraction: amino acid sequence of S
A;Reference number: A92972; MUID:85291729; PMID:4031854
A;Accession: A92972
A;Molecule type: protein
A;Residues: 2-92 <JEN>
R;Baudier, J.; Glasser, N.; Haglid, K.; Gerard, D.
Biochim. Biophys. Acta 790, 164-173, 1984
A;Title: Purification, characterization and ion binding properties of human brain S100b
A;Reference number: A90653; MUID:85023393; PMID:6487634
A;Contents: annotation; metal ion-binding properties
C;Comment: This protein binds p53, tubulin and many other proteins at physiological con
C;Comment: S-100 is an intracellular protein that binds calcium. It binds zinc more tig
different affinities exist for both ions on each monomer. Physiological concentrations
nding sites.
C;Comment: This protein is expressed predominantly in brain tissue by astroglial cells.
C;Comment: The homodimer contains disulfide bonds, but the bond pattern has not been de
C;Genetics:
A;Gene: GDB:S100B
A;Cross-references: GDB:120360; OMIM:176990
A;Map position: 21q22.3-21q22.3
A;Introns: 46/3
A;Note: the first intron occurs before the initiator codon
C;Complex: homodimer; heterodimer with S-100 protein alpha chain (see PIR:BCHUIA)
C;Superfamily: S-100 protein; calmodulin repeat homology
C;Keywords: blocked amino end; brain; calcium binding; EF hand; heterodimer; homodimer;
F;2-92/Product: S-100 protein beta chain #status experimental <MAT>
F;6-40/Domain: calmodulin repeat homology <EF1>
F;49-81/Domain: calmodulin repeat homology <EF2>
F;2/Modified site: blocked amino end (Ser) (in mature form) (probably acetylated) #stat
F;19,22,24,27,32/Binding site: calcium (Ser, Glu, Asp, Lys, Glu) #status predicted
F;62,64,66,68,73/Binding site: calcium (Asp, Asp, Asp, Glu, Glu) #status predicted

Query Match 34.7%; Score 162; DB 1; Length 92;
Best Local Similarity 37.4%; Pred. No. 1.3e-08;
Matches 34; Conservative 18; Mismatches 39; Indels 0; Gaps 0;

```
Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ::||  :  :|::|||||| | |    | | |||:||  || : |:  |:|  :||: : |
Db          2 SELEKAMVALIDVFHQYSGREGDKHKLKKSELKELINNELSHFLEEIKEQEVVDKVMETL 61

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              | : | :  |:||:  |  |    |:    |
Db         62 DNDGDGECDFQEFMAFVAMVTTACHEFFEHE 92
```

RESULT 15
BCBOIA
S-100 protein alpha chain - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Nov-1983 #sequence_revision 06-Feb-1995 #text_change 24-Nov-1999
C;Accession: A24156; A91110; S54346; A03078
R;Kuwano, R.; Maeda, T.; Usui, H.; Araki, K.; Yamakuni, T.; Ohshima, Y.; Kurihara, T.;
FEBS Lett. 202, 97-101, 1986
A;Title: Molecular cloning of cDNA of S100alpha subunit mRNA.
A;Reference number: A24156; MUID:86248083; PMID:3755105
A;Accession: A24156
A;Molecule type: mRNA
A;Residues: 1-94 <KUW>
R;Isobe, T.; Okuyama, T.
Eur. J. Biochem. 116, 79-86, 1981
A;Title: The amino-acid sequence of the alpha subunit in bovine brain S-100 a protein.
A;Reference number: A91110; MUID:81236562; PMID:7250124
A;Accession: A91110
A;Molecule type: protein
A;Residues: 2-64,'D',66-94 <ISO>
R;Baudier, J.; Gerard, D.
Biochemistry 22, 3360-3369, 1983
A;Title: Ions binding to S100 proteins: structural changes induced by calcium and zinc
A;Reference number: A90471; MUID:84000339; PMID:6615778
A;Contents: annotation; metal ion-binding properties
R;Okazaki, K.; Obata, N.H.; Inoue, S.; Hidaka, H.
Biochem. J. 306, 551-555, 1995
A;Title: S100-beta is a target protein of neurocalcin delta, an abundant isoform in gli
A;Reference number: S54343; MUID:95194333; PMID:7887910
A;Accession: S54346
A;Molecule type: protein
A;Residues: 24-33 <OKA>
C;Comment: The S-100 protein is composed of two related polypeptide chains, alpha and b
brain proteins, S-100 is also found in a variety of other tissues.
C;Comment: S-100 is an intracellular protein that weakly binds calcium. It binds zinc v
es, with different affinities, exist for both ions on each monomer. Physiological concer
alcium-binding sites.

```
';Superfamily: S-100 protein; calmodulin repeat homology
';Keywords: blocked amino end; brain; calcium binding; EF hand; zinc
';2-94/Product: S-100 protein alpha chain #status predicted <MAT>
';7-41/Domain: calmodulin repeat homology <EF1>
';50-82/Domain: calmodulin repeat homology <EF2>
';2/Modified site: blocked amino end (Gly) (in mature form) (probably acetylated) #statu
';20,23,25,28,33/Binding site: calcium (Ser, Glu, Asp, Lys, Glu) #status predicted
';63,65,67,69,74/Binding site: calcium (Asp, Asn, Asp, Glu, Glu) #status predicted

  Query Match             34.5%;  Score 161;  DB 1;  Length 94;
  Best Local Similarity   38.4%;  Pred. No. 1.6e-08;
  Matches   33;  Conservative   20;  Mismatches   33;  Indels    0;  Gaps    0;

)y          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ::||  :| :||:|| :| : |    | |:|||:|:  ||   |   ||   :||: : |
)b          3 SELETAMETLINVFHAHSGKEGDKYKLSKKELKELLQTELSGFLDAQKDADAVDKVMKEL 62

)y         61 DANQDEQVSFKEFVVLVTDVLITAHD 86
              | | | :| |:|:||||  : :  ::
)b         63 DENGDGEVDFQEYVVLVAALTVACNN 88


!earch completed: March  5, 2004, 09:49:26
'ob time : 21 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

'M protein - protein search, using sw model

'un on: March 5, 2004, 09:36:18 ; Search time 14 Seconds
(without alignments)
338.456 Million cell updates/sec

'itle: US-09-646-651C-1
'erfect score: 467
'equence: 1 TKLEDHLEGIINIFHQYSVR.....EFVVLVTDVLITAHDNIHKE 91

'coring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

'earched: 141681 seqs, 52070155 residues

'otal number of hits satisfying chosen parameters: 141681

'inimum DB seq length: 0
'aximum DB seq length: 2000000000

'ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

'atabase : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

esult No.	Score	% Query Match	Length	DB	ID	Description
1	467	100.0	91	1	S112_PIG	P80310 sus scrofa
2	371	79.4	91	1	S112_BOVIN	P79105 bos taurus
3	332	71.1	91	1	S112_HUMAN	P80511 homo sapien
4	302	64.7	81	1	S112_RABIT	O77791 oryctolagus
5	263.5	56.4	122	1	S109_BOVIN	P28783 bos taurus
6	220	47.1	119	1	M126_CHICK	P28318 gallus gall
7	216.5	46.4	118	1	S109_RABIT	P50117 oryctolagus
8	213.5	45.7	114	1	S109_HUMAN	P06702 homo sapien
9	204.5	43.8	112	1	S109_RAT	P50116 rattus norv
10	181	38.8	95	1	S10P_HUMAN	P25815 homo sapien
11	178.5	38.2	112	1	S109_MOUSE	P31725 mus musculu
12	166	35.5	92	1	S10I_ICTPU	Q91061 ictalurus p
13	164	35.1	91	1	S10B_RAT	P04631 rattus norv
14	164	35.1	98	1	S10Z_HUMAN	Q8wxg8 homo sapien
15	163	34.9	88	1	S108_MOUSE	P27005 mus musculu
16	163	34.9	88	1	S108_RAT	P50115 rattus norv
17	163	34.9	91	1	S10B_BOVIN	P02638 bos taurus
18	163	34.9	91	1	S10B_MOUSE	P50114 mus musculu
19	163	34.9	93	1	S10A_RAT	P35467 rattus norv
20	163	34.9	101	1	S104_MOUSE	P07091 mus musculu
21	162	34.7	91	1	S10B_HUMAN	P04271 homo sapien
22	162	34.7	93	1	S10A_MOUSE	P56565 mus musculu
23	161	34.5	93	1	S10A_BOVIN	P02639 bos taurus
24	160	34.3	93	1	S10A_HUMAN	P23297 homo sapien
25	160	34.3	101	1	S104_RAT	P05942 rattus norv
26	156	33.4	101	1	S104_HUMAN	P26447 homo sapien
27	154	33.0	102	1	S111_RABIT	P24480 oryctolagus
28	153	32.8	100	1	S104_BOVIN	P35466 bos taurus
29	151	32.3	93	1	S108_HUMAN	P05109 homo sapien
30	150	32.1	105	1	S111_HUMAN	P31949 homo sapien
31	147	31.5	93	1	S105_MOUSE	O88945 mus musculu
32	147	31.5	101	1	S111_CHICK	P24479 gallus gall
33	146	31.3	98	1	S111_MOUSE	P50543 mus musculu
34	144.5	30.9	78	1	S10D_PIG	P02632 sus scrofa
35	143.5	30.7	92	1	S105_HUMAN	P33763 homo sapien
36	143	30.6	97	1	S102_HUMAN	P29034 homo sapien
37	142.5	30.5	101	1	S107_BOVIN	Q28050 bos taurus
38	140.5	30.1	98	1	S113_HUMAN	Q99584 homo sapien
39	139.5	29.9	78	1	S10D_HUMAN	P29377 homo sapien
40	138	29.6	97	1	S102_BOVIN	P10462 bos taurus
41	136	29.1	90	1	S106_HUMAN	P06703 homo sapien
42	135.5	29.0	96	1	S110_MOUSE	P08207 mus musculu
43	135	28.9	78	1	S10D_BOVIN	P02633 bos taurus
44	135	28.9	90	1	S106_RABIT	P30801 oryctolagus
45	134.5	28.8	94	1	S110_RAT	P05943 rattus norv

ALIGNMENTS

RESULT 1
S112_PIG
ID S112_PIG STANDARD; PRT; 91 AA.
AC P80310;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calgranulin C (CAGC).
GN S100A12.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE.
RC TISSUE=Granulocyte;
RX MEDLINE=95050708; PubMed=7961855;
RA Dell'Angelica E.C., Schleicher C.H., Santome J.A.;
RT "Primary structure and binding properties of calgranulin C, a novel
RT S100-like calcium-binding protein from pig granulocytes.";
RL J. Biol. Chem. 269:28929-28936(1994).
CC -!- TISSUE SPECIFICITY: Found essentially in granulocytes with small
CC amounts found in lymphocytes.
CC -!- MISCELLANEOUS: In the absence of zinc binds one calcium ion per
CC molecule, in the presence of zinc binds two calcium ions per
CC molecule.
CC -!- SIMILARITY: Belongs to the S-100 family.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
DR PIR; A55406; A55406.
DR HSSP; P80511; 1E8A.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; efhand; 1.
DR Pfam; PF01023; S_100; 1.
DR ProDom; PD003407; CaBP_S100; 1.
DR ProDom; PD000012; EF-hand; 1.
DR PROSITE; PS00018; EF_HAND; FALSE_NEG.
DR PROSITE; PS00303; S100_CABP; 1.
KW Calcium-binding; Zinc; Metal-binding.
FT CA_BIND 18 31 EF-HAND 1 (LOW AFFINITY) (BY SIMILARITY).
FT CA_BIND 61 72 EF-HAND 2 (HIGH AFFINITY) (BY
FT SIMILARITY).
SQ SEQUENCE 91 AA; 10614 MW; B4204461432D7FCE CRC64;

Query Match 100.0%; Score 467; DB 1; Length 91;
Best Local Similarity 100.0%; Pred. No. 9.5e-37;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              |||||||||||||||||||||||||||||||
Db         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
```

```
RESULT 2
S112_BOVIN
ID   S112_BOVIN     STANDARD;      PRT;    91 AA.
AC   P79105;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Calgranulin C (CAGC) (Calcium-binding protein in amniotic fluid 1)
DE   (CAAF1) (RAGE binding protein).
GN   S100A12 OR CAAF1.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Oesophagus;
RX   MEDLINE=96298783; PubMed=8718672;
RA   Hitomi J., Yamaguchi K., Kikuchi Y., Kimura T., Maruyama K.,
RA   Nagasaki K.;
RT   "A novel calcium-binding protein in amniotic fluid, CAAF1: its
RT   molecular cloning and tissue distribution.";
RL   J. Cell Sci. 109:805-815(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=99325504; PubMed=10399917;
RA   Hofmann M.A., Drury S., Fu C., Qu W., Taguchi A., Lu Y., Avila C.,
RA   Kambham N., Bierhaus A., Nawroth P., Neurath M.F., Slattery T.,
RA   Beach D., McClary J., Nagashima M., Morser J., Stern D.,
RA   Schmidt A.M.;
RT   "RAGE mediates a novel proinflammatory axis: a central cell surface
RT   receptor for S100/calgranulin polypeptides.";
RL   Cell 97:889-901(1999).
CC   -!- SIMILARITY: Belongs to the S-100 family.
CC   -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D49548; BAA08496.1; -.
DR   EMBL; AF011757; AAB65423.1; -.
DR   HSSP; P80511; 1E8A.
DR   InterPro; IPR001751; CaBP_S100.
DR   InterPro; IPR002048; EF-hand.
DR   Pfam; PF00036; efhand; 1.
DR   Pfam; PF01023; S_100; 1.
DR   ProDom; PD003407; CaBP_S100; 1.
DR   ProDom; PD000012; EF-hand; 1.
DR   PROSITE; PS00018; EF_HAND; 1.
DR   PROSITE; PS00303; S100_CABP; 1.
KW   Calcium-binding; Zinc; Metal-binding.
FT   INIT_MET      0      0       BY SIMILARITY.
FT   CA_BIND      18     31       EF-HAND 1 (LOW AFFINITY) (BY SIMILARITY).
FT   CA_BIND      61     72       EF-HAND 2 (HIGH AFFINITY) (BY
FT                                SIMILARITY).
SQ   SEQUENCE   91 AA;  10554 MW;  66FBC3C1B0354482 CRC64;

  Query Match             79.4%;  Score 371;  DB 1;  Length 91;
  Best Local Similarity   81.3%;  Pred. No. 7.1e-28;
  Matches   74;  Conservative    9;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              |||||||||||||||||||:||:||| |||||||||||||| ||:||||| |||||||:|
Db          1 TKLEDHLEGIINIFHQYSVRVGHFDTLNKRELKQLITKELPKTLQNTKDQPTIDKIFQDL 60

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              ||::|  |||:||||||: || ||| :||||
Db         61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHKE 91


RESULT 3
S112_HUMAN
ID   S112_HUMAN     STANDARD;      PRT;    91 AA.
AC   P80511; P83219;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Calgranulin C (CAGC) (CGRP) (Neutrophil S100 protein) (Calcium-binding
DE   protein in amniotic fluid 1) (CAAF1) (p6) [Contains: Calcitermin].
GN   S100A12.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97138564; PubMed=8985590;
RA   Wicki R., Marenholz I., Mischke D., Schaefer B.W., Heizmann C.W.;
RT   "Characterization of the human S100A12 (calgranulin C, p6, CAAF1,
RT   CGRP) gene, a new member of the S100 gene cluster on chromosome
RT   1q21.";
RL   Cell Calcium 20:459-464(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96192053; PubMed=8619860;
RA   Yamamura T., Hitomi J., Nagasaki K., Suzuki M., Takahashi E.,
RA   Saito S., Tsukada T., Yamaguchi K.;
RT   "Human CAAF1 gene -- molecular cloning, gene structure, and chromosome
RT   mapping.";
RL   Biochem. Biophys. Res. Commun. 221:356-360(1996).
RN   [3]
RP   SEQUENCE.
RX   MEDLINE=96192069; PubMed=8619876;
RA   Marti T., Erttmann K.D., Gallin M.Y.;
RT   "Host-parasite interaction in human onchocerciasis: identification
RT   and sequence analysis of a novel human calgranulin.";
RL   Biochem. Biophys. Res. Commun. 221:454-458(1996).
RN   [4]
RP   SEQUENCE.
RC   TISSUE=Neutrophils;
RX   MEDLINE=96332419; PubMed=8769108;
RA   Ilg E.C., Troxler H., Buergisser D.M., Kuster T., Markert M.,
RA   Guignard F., Hunziker P., Birchler N., Heizmann C.W.;
RT   "Amino acid sequence determination of human S100A12 (P6, calgranulin
RT   C, CGRP, CAAF1) by tandem mass spectrometry.";
RL   Biochem. Biophys. Res. Commun. 225:146-150(1996).
RN   [5]
RP   SEQUENCE OF 1-20.
RX   MEDLINE=95351965; PubMed=7626002;
RA   Guignard F., Mauel J., Markert M.;
RT   "Identification and characterization of a novel human neutrophil
RT   protein related to the S100 family.";
RL   Biochem. J. 309:395-401(1995).
RN   [6]
RP   SEQUENCE OF 77-91, ANTIMICROBIAL ACTIVITY, AND MASS SPECTROMETRY.
RC   TISSUE=Nasal mucus;
RX   MEDLINE=21413725; PubMed=11522286;
RA   Cole A.M., Kim Y.-H., Tahk S., Hong T., Weis P., Waring A.J., Ganz T.;
RT   "Calcitermin, a novel antimicrobial peptide isolated from human airway
RT   secretions.";
RL   FEBS Lett. 504:5-10(2001).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).
RX   MEDLINE=21065388; PubMed=11134923;
RA   Moroz O.V., Antson A.A., Murshudov G.N., Maitland N.J., Dodson G.G.,
RA   Wilson K.S., Skibshoj I., Lukanidin E.M., Bronstein I.B.;
RT   "The three-dimensional structure of human S100A12.";
RL   Acta Crystallogr. D 57:20-29(2001).
```

```
:C   -!- FUNCTION: Calcitermin possesses antifungal activity against
:C       C.albicans and is also active against E.coli and P.aeruginosa but
:C       not L.monocytogenes and S.aureus.
:C   -!- SUBUNIT: Homodimer.
:C   -!- TISSUE SPECIFICITY: Monocytes and lymphocytes.
:C   -!- MASS SPECTROMETRY: MW=10444; METHOD=Electrospray; RANGE=1-91.
:C   -!- MASS SPECTROMETRY: MW=1688.9; METHOD=MALDI; RANGE=77-91.
:C   -!- SIMILARITY: Belongs to the S-100 family.
:C   -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
:C   ---------------------------------------------------------------------------
:C   This SWISS-PROT entry is copyright. It is produced through a collaboration
:C   between the Swiss Institute of Bioinformatics and the EMBL outstation -
:C   the European Bioinformatics Institute. There are no restrictions on its
:C   use by non-profit institutions as long as its content is in no way
:C   modified and this statement is not removed. Usage by and for commercial
:C   entities requires a license agreement (See http://www.isb-sib.ch/announce/
:C   or send an email to license@isb-sib.ch).
:C   ---------------------------------------------------------------------------
)R   EMBL; X97859; CAA66453.1; -.
)R   EMBL; X98288; CAA66934.1; -.
)R   EMBL; X98289; CAA66934.1; JOINED.
)R   EMBL; X98290; CAA66934.1; JOINED.
)R   EMBL; X98289; CAB94792.1; -.
)R   EMBL; X98290; CAB94792.1; JOINED.
)R   EMBL; D49549; BAA08497.1; -.
)R   EMBL; D83664; BAA12036.1; -.
)R   EMBL; D83657; BAA12030.1; -.
)R   PIR; JC4712; JC4712.
)R   PDB; 1E8A; 08-JAN-01.
)R   PDB; 1GQM; 07-MAR-02.
)R   Genew; HGNC:10489; S100A12.
)R   MIM; 603112; -.
)R   GO; GO:0005829; C:cytosol; TAS.
)R   GO; GO:0005626; C:insoluble fraction; TAS.
)R   GO; GO:0005509; F:calcium ion binding; TAS.
)R   GO; GO:0006954; P:inflammatory response; TAS.
)R   InterPro; IPR001751; CaBP_S100.
)R   InterPro; IPR002048; EF-hand.
)R   Pfam; PF00036; efhand; 1.
)R   Pfam; PF01023; S_100; 1.
)R   ProDom; PD003407; CaBP_S100; 1.
)R   ProDom; PD000012; EF-hand; 1.
)R   PROSITE; PS00018; EF_HAND; FALSE_NEG.
)R   PROSITE; PS00303; S100_CABP; 1.
:W   Calcium-binding; Zinc; Metal-binding; Antibiotic; Fungicide;
:W   3D-structure.
'T   INIT_MET      0      0
'T   PEPTIDE      77     91       CALCITERMIN.
'T   CA_BIND      18     31       EF-HAND 1 (LOW AFFINITY) (BY SIMILARITY).
'T   CA_BIND      61     72       EF-HAND 2 (HIGH AFFINITY) (BY
'T                                SIMILARITY).
'T   HELIX         2     18
'T   TURN         19     19
'T   TURN         24     25
'T   STRAND       26     27
'T   HELIX        29     39
'T   TURN         41     43
'T   TURN         45     48
'T   HELIX        50     60
'T   TURN         62     63
'T   STRAND       68     69
'T   HELIX        70     85
:Q   SEQUENCE   91 AA;  10444 MW;  325685EA8695F6B7 CRC64;

  Query Match             71.1%;  Score 332;  DB 1;  Length 91;
  Best Local Similarity   70.3%;  Pred. No. 2.9e-24;
  Matches   64;  Conservative   10;  Mismatches   17;  Indels    0;  Gaps    0;

)y          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ||||:|||||:|||||||||| ||:||| | |||||:|||| ||:|| ||:  ||:||| |
)b          1 TKLEEHLEGIVNIFHQYSVRKGHFDTLSKGELKQLLTKELANTIKNIKDKAVIDEIFQGL 60

)y         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
                ||||||| |:||: ||   |  || : |||
Db         61 DANQDEQVDFQEFISLVAIALKAAHYHTHKE 91


RESULT 4
S112_RABIT
ID   S112_RABIT     STANDARD;      PRT;    81 AA.
AC   O77791;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Calgranulin C (CAGC) (Fragment).
GN   S100A12.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   STRAIN=New Zealand white; TISSUE=Neutrophils;
RX   MEDLINE=96355278; PubMed=8702688;
RA   Yang Z., Deveer M.J., Gardiner E.E., Devenish R.J., Handley C.J.,
RA   Underwood J.R., Robinson H.C.;
RT   "Rabbit polymorphonuclear neutrophils form 35S-labeled S-sulfo-
RT   calgranulin C when incubated with inorganic [35S]sulfate.";
RL   J. Biol. Chem. 271:19802-19809(1996).
CC   -!- SIMILARITY: Belongs to the S-100 family.
CC   -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF091848; AAC61770.1; -.
DR   HSSP; P80511; 1E8A.
DR   InterPro; IPR001751; CaBP_S100.
DR   InterPro; IPR002048; EF-hand.
DR   Pfam; PF00036; efhand; 1.
DR   Pfam; PF01023; S_100; 1.
DR   ProDom; PD003407; CaBP_S100; 1.
DR   ProDom; PD000012; EF-hand; 1.
DR   PROSITE; PS00018; EF_HAND; 1.
DR   PROSITE; PS00303; S100_CABP; 1.
KW   Calcium-binding.
FT   NON_TER       1      1
FT   CA_BIND       8     21       EF-HAND 1 (LOW AFFINITY) (BY SIMILARITY).
FT   CA_BIND      51     62       EF-HAND 2 (HIGH AFFINITY) (BY
FT                                SIMILARITY).
SQ   SEQUENCE   81 AA;  9401 MW;  95E67A209180CB66 CRC64;

  Query Match             64.7%;  Score 302;  DB 1;  Length 81;
  Best Local Similarity   70.4%;  Pred. No. 1.5e-21;
  Matches   57;  Conservative   11;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         11 INIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNLDANQDEQVSF 70
              |||||||||| |||||| | |||:||| || ||:|||||| |:|:||::|| : | || |
Db          1 INIFHQYSVRTGHYDTLSKCELKKLITTELVNTIKNTKDQATVDRIFRDLDEDGDHQVDF 60

Qy         71 KEFVVLVTDVLITAHDNIHKE 91
              |||: |:  ||:|||:|||||
Db         61 KEFLSLLASVLVTAHENIHKE 81


RESULT 5
S109_BOVIN
ID   S109_BOVIN     STANDARD;      PRT;   122 AA.
AC   P28783;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
```

```
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Calgranulin B (Neutrophil cytosolic 23 kDa protein) (P23) (BEE22)
DE   (Fragment).
GN   S100A9.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Oesophageal epithelium;
RX   MEDLINE=93280230; PubMed=8505358;
RA   Tang T.K., Hong T.-M., Lin C.-Y., Lai M.-L., Liu C.H.L., Lo H.-J.,
RA   Wang M.-E., Chen L.B., Chen W.-T., Ip W., Lin D.C., Lin J.J.-C.,
RA   Lin S., Sun T.-T., Wang E., Wang J.L., Wu R., Wu C.-W., Chien S.;
RT   "Nuclear proteins of the bovine esophageal epithelium. I. Monoclonal
RT   antibody W2 specifically reacts with condensed nuclei of
RT   differentiated superficial cells.";
RL   J. Cell Sci. 104:237-247(1993).
RN   [2]
RP   SEQUENCE OF 4-56.
RC   TISSUE=Neutrophils;
RX   MEDLINE=92304974; PubMed=1610833;
RA   Dianoux A.-C., Stasia M.-J., Garin J., Gagnon J., Vignais P.V.;
RT   "The 23-kilodalton protein, a substrate of protein kinase C, in
RT   bovine neutrophil cytosol is a member of the S100 family.";
RL   Biochemistry 31:5898-5905(1992).
CC   -!- SUBUNIT: Disulfide linked heterodimer of a 7/11 kDa and a 22/23
CC       kDa subunits.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; loosely associated to the
CC       cytoskeleton.
CC   -!- TISSUE SPECIFICITY: Found essentially in phagocytic cells.
CC   -!- PTM: Phosphorylated by protein kinase C.
CC   -!- SIMILARITY: Belongs to the S-100 family.
CC   -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
DR   HSSP; P80511; 1E8A.
DR   InterPro; IPR001751; CaBP_S100.
DR   InterPro; IPR002048; EF-hand.
DR   Pfam; PF00036; efhand; 1.
DR   Pfam; PF01023; S_100; 1.
DR   ProDom; PD003407; CaBP_S100; 1.
DR   ProDom; PD000012; EF-hand; 1.
DR   PROSITE; PS00018; EF_HAND; PARTIAL.
DR   PROSITE; PS00303; S100_CABP; 1.
KW   Calcium-binding; Phosphorylation.
FT   NON_TER       1      1
FT   CA_BIND      19     32       EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT   CA_BIND      63     74       EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
SQ   SEQUENCE   122 AA;  13673 MW;  F3CA8C48806BECCD CRC64;

  Query Match             56.4%;  Score 263.5;  DB 1;  Length 122;
  Best Local Similarity   53.3%;  Pred. No. 8.4e-18;
  Matches   48;  Conservative   23;  Mismatches   18;  Indels    1;  Gaps    1;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTL-KNTKDQGTIDKIFQN 59
              :::|  :| ||||||||||||||||||::| |||: ||||| | |  |::  |::| ::
Db          2 SQMESSIETIINIFHQYSVRLGHYDTLIQKEFKQLVQKELPNFLKKQKKNEAAINEIMED 61

Qy         60 LDANQDEQVSFKEFVVLVTDVLITAHDNIH 89
              || | |:|:||:||::||  : : :|: :|
Db         62 LDTNVDKQLSFEEFIMLVARLTVASHEEMH 91


RESULT 6
M126_CHICK
ID   M126_CHICK     STANDARD;      PRT;   119 AA.
AC   P28318;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Protein MRP-126.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=White leghorn; TISSUE=Bone marrow;
RX   MEDLINE=92195690; PubMed=1549365;
RA   Nakano T., Graf T.;
RT   "Identification of genes differentially expressed in two types of
RT   v-myb-transformed avian myelomonocytic cells.";
RL   Oncogene 7:527-534(1992).
CC   -!- TISSUE SPECIFICITY: Expressed in v-myb-transformed myelomonocytic
CC       cells.
CC   -!- SIMILARITY: Belongs to the S-100 family.
CC   -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X61200; -; NOT_ANNOTATED_CDS.
DR   HSSP; P04271; 1UWO.
DR   InterPro; IPR001751; CaBP_S100.
DR   InterPro; IPR002048; EF-hand.
DR   Pfam; PF00036; efhand; 1.
DR   Pfam; PF01023; S_100; 1.
DR   ProDom; PD003407; CaBP_S100; 1.
DR   ProDom; PD000012; EF-hand; 1.
DR   PROSITE; PS00018; EF_HAND; 1.
DR   PROSITE; PS00303; S100_CABP; 1.
KW   Calcium-binding.
FT   CA_BIND      29     42       EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT   CA_BIND      72     83       EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
SQ   SEQUENCE   119 AA;  14065 MW;  2D268DAF6309AD7A CRC64;

  Query Match             47.1%;  Score 220;  DB 1;  Length 119;
  Best Local Similarity   47.2%;  Pred. No. 8.6e-14;
  Matches   42;  Conservative   25;  Mismatches   22;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ::||  :: ||::|||||  | |  ||| ::||| || |:| | ||: |:| :||:|||::|
Db         12 SELEKAIDVIIDVFHQYSRREGDKDTLTRKELKLLIEKQLANYLKHVKNQVSIDQIFKDL 71

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIH 89
              | |:|:|:|| | ::|:  | :  |:::|
Db         72 DNNKDQQLSFGEVMLLIIRVTVATHEHLH 100


RESULT 7
S109_RABIT
ID   S109_RABIT     STANDARD;      PRT;   118 AA.
AC   P50117;
DT   01-OCT-1996 (Rel. 34, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Calgranulin B (Migration inhibitory factor-related protein 14)
DE   (MRP-14) (Fragment).
GN   S100A9 OR MRP-14.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=New Zealand white; TISSUE=Neutrophils;
RX   MEDLINE=96355278; PubMed=8702688;
RA   Yang Z., de Veer M.J., Gardiner E.E., Devenish R.J., Handley C.J.,
RA   Underwood J.R., Robinson H.C.;
```

```
RT   "Rabbit polymorphonuclear neutrophils form 35S-labeled S-sulfo-
RT   calgranulin C when incubated with inorganic [35S]sulfate.";
RL   J. Biol. Chem. 271:19802-19809(1996).
RN   [2]
RP   SEQUENCE OF 45-82 FROM N.A.
RC   STRAIN=New Zealand white;
RX   MEDLINE=94198229; PubMed=8148323;
RA   Mori S., Goto K., Goto F., Mutakami K., Ohkawara S., Yoshinaga M.;
RT   "Dynamic changes in mRNA expression of neutrophils during the course
RT   of acute inflammation in rabbits.";
RL   Int. Immunol. 6:149-156(1994).
CC   -!- SIMILARITY: Belongs to the S-100 family.
CC   -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF091849; AAC61771.1; -.
DR   EMBL; D17404; BAA04227.1; -.
DR   PIR; I46861; I46861.
DR   HSSP; P80511; 1E8A.
DR   InterPro; IPR001751; CaBP_S100.
DR   InterPro; IPR002048; EF-hand.
DR   Pfam; PF00036; efhand; 1.
DR   Pfam; PF01023; S_100; 1.
DR   ProDom; PD003407; CaBP_S100; 1.
DR   ProDom; PD000012; EF-hand; 1.
DR   PROSITE; PS00018; EF_HAND; 1.
DR   PROSITE; PS00303; S100_CABP; 1.
KW   Calcium-binding; Repeat.
FT   NON_TER       1      1
FT   CA_BIND       9     22       EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT   CA_BIND      53     64       EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
FT   DOMAIN      103    118       2 X 8 AA TANDEM REPEATS OF G-H-G-H-
FT                                G-H-S-H.
FT   REPEAT      103    110       1.
FT   REPEAT      111    118       2.
SQ   SEQUENCE   118 AA;  13292 MW;  7496118E21AD5C41 CRC64;

  Query Match             46.4%;  Score 216.5;  DB 1;  Length 118;
  Best Local Similarity   51.2%;  Pred. No. 1.8e-13;
  Matches   42;  Conservative   20;  Mismatches   19;  Indels    1;  Gaps    1;

Qy         10 IINIFHQYSVRLGHYDTLIKRELKQLITKELPNTL-KNTKDQGTIDKIFQNLDANQDEQV 68
              ||||||||||:|  |:| ::| |||: ||| | | |  :|:  |: | ::|| |||:|:
Db          1 IINIFHQYSVRVGPRDSLSQKEFKQLVQKELHNFLKKEARDEKAINDIMEDLDTNQDKQL 60

Qy         69 SFKEFVVLVTDVLITAHDNIHK 90
              ||:|||:|:  ::  :|: :||
Db         61 SFEEFVILMARLVHASHEEMHK 82


RESULT 8
S109_HUMAN
ID   S109_HUMAN     STANDARD;      PRT;   114 AA.
AC   P06702; Q9NYM0; Q9UCJ1;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Calgranulin B (Migration inhibitory factor-related protein 14)
DE   (MRP-14) (P14) (Leukocyte L1 complex heavy chain) (S100 calcium-
DE   binding protein A9) (Calprotectin L1H subunit).
GN   S100A9 OR MRP14 OR CAGB.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88039099; PubMed=3313057;
RA   Odink K., Cerletti N., Bruggen J., Clerc R.G., Tarcsay L., Zwaldo G.,
RA   Gerhards G., Schlegel R., Sorg C.;
RT   "Two calcium-binding proteins in infiltrate macrophages of rheumatoid
RT   arthritis.";
RL   Nature 330:80-82(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88302148; PubMed=3405210;
RA   Lagasse E., Clerc R.G.;
RT   "Cloning and expression of two human genes encoding calcium-binding
RT   proteins that are regulated during myeloid differentiation.";
RL   Mol. Cell. Biol. 8:2402-2410(1988).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89255276; PubMed=2656677;
RA   Murao S., Collart F.R., Huberman E.;
RT   "A protein containing the cystic fibrosis antigen is an inhibitor of
RT   protein kinases.";
RL   J. Biol. Chem. 264:8356-8360(1989).
RN   [4]
RP   SEQUENCE FROM N.A., AND VARIANT ARG-20.
RA   Wang M., Xu X., Cai Y., Xu H., Han Y., Xu Z., Wu M.;
RT   "Human gene for migration inhibitory factor-related protein 14
RT   (MRP14), variant allele.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [6]
RP   SEQUENCE OF 84-114, AND PHOSPHORYLATION.
RX   MEDLINE=90044075; PubMed=2478889;
RA   Edgeworth J., Freemont P., Hogg N.;
RT   "Ionomycin-regulated phosphorylation of the myeloid calcium-binding
RT   protein p14.";
RL   Nature 342:189-192(1989).
RN   [7]
RP   SEQUENCE OF 11-19; 26-37 AND 94-107.
RC   TISSUE=Keratinocytes;
RX   MEDLINE=93162043; PubMed=1286667;
RA   Rasmussen H.H., van Damme J., Puype M., Gesser B., Celis J.E.,
RA   Vandekerckhove J.;
RT   "Microsequences of 145 proteins recorded in the two-dimensional gel
RT   protein database of normal human epidermal keratinocytes.";
RL   Electrophoresis 13:960-969(1992).
RN   [8]
RP   SEQUENCE OF 5-34.
RX   MEDLINE=93139333; PubMed=8423249;
RA   Miyasaki K.T., Bodeau A.L., Murthy A.R., Lehrer R.I.;
RT   "In vitro antimicrobial activity of the human neutrophil cytosolic S-
RT   100 protein complex, calprotectin, against Capnocytophaga sputigena.";
```

```
RL   J. Dent. Res. 72:517-523(1993).
CC   -!- FUNCTION: Expressed by macrophages in acutely inflammated tissues
CC       and in chronic inflammations. Seem to be an inhibitor of protein
CC       kinases. Also expressed in epithelial cells constitutively or
CC       induced during dermatoses. May interact with components of the
CC       intermediate filaments in monocytes and epithelial cells.
CC   -!- MISCELLANEOUS: Has been shown to bind calcium.
CC   -!- SIMILARITY: Belongs to the S-100 family.
CC   -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC   ---------------------------------------------------------------------
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X06233; CAA29579.1; -.
DR   EMBL; M21064; AAA36326.1; -.
DR   EMBL; M26311; AAA68480.1; -.
DR   EMBL; AF237581; AAF62536.1; -.
DR   EMBL; AF237582; AAF62537.1; -.
DR   EMBL; A12029; CAA01002.1; -.
DR   EMBL; A12032; CAA01004.1; -.
DR   EMBL; BC047681; AAH47681.1; -.
DR   PIR; B31848; B31848.
DR   PDB; 1IRJ; 14-JAN-03.
DR   SWISS-2DPAGE; P06702; HUMAN.
DR   Aarhus/Ghent-2DPAGE; 5007; IEF.
DR   Aarhus/Ghent-2DPAGE; 6010; IEF.
DR   Aarhus/Ghent-2DPAGE; 6017; IEF.
DR   Aarhus/Ghent-2DPAGE; 7013; IEF.
DR   PMMA-2DPAGE; P06702; -.
DR   Genew; HGNC:10499; S100A9.
DR   MIM; 123886; -.
DR   GO; GO:0005509; F:calcium ion binding; TAS.
DR   GO; GO:0004871; F:signal transducer activity; TAS.
DR   GO; GO:0007267; P:cell-cell signaling; TAS.
DR   GO; GO:0006954; P:inflammatory response; TAS.
DR   InterPro; IPR001751; CaBP_S100.
DR   InterPro; IPR002048; EF-hand.
DR   Pfam; PF00036; efhand; 1.
DR   Pfam; PF01023; S_100; 1.
DR   ProDom; PD003407; CaBP_S100; 1.
DR   PROSITE; PS00018; EF_HAND; 1.
DR   PROSITE; PS00303; S100_CABP; 1.
KW   Calcium-binding; Macrophage; Phosphorylation; Polymorphism;
KW   3D-structure.
FT   CA_BIND      23     36       EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT   CA_BIND      67     78       EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
FT   MOD_RES     113    113       PHOSPHORYLATION.
FT   VARIANT      20     20       H -> R.
FT                                /FTId=VAR_013008.
FT   CONFLICT      6      6       S -> H (IN REF. 8).
FT   CONFLICT     25     25       K -> F (IN REF. 8).
FT   CONFLICT     28     28       H -> L (IN REF. 8).
SQ   SEQUENCE   114 AA;  13242 MW;  C3BE19729E14C078 CRC64;

  Query Match             45.7%;  Score 213.5;  DB 1;  Length 114;
  Best Local Similarity   45.1%;  Pred. No. 3.3e-13;
  Matches   41;  Conservative   25;  Mismatches   24;  Indels    1;  Gaps    1;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTL-KNTKDQGTIDKIFQN 59
              ::|| ::| ||| |||||||||||| ||| : | |:|: |:| | | | |::  |: | ::
Db          6 SQLERNIETIINTFHQYSVKLGHPDTLNQGEFKELVRKDLQNFLKKENKNEKVIEHIMED 65

Qy         60 LDANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
              || | |:|:||:||::|:  :   :|: :|:
Db         66 LDTNADKQLSFEEFIMLMARLTWASHEKMHE 96

RESULT 9
S109_RAT
ID   S109_RAT       STANDARD;      PRT;   112 AA.
AC   P50116;
DT   01-OCT-1996 (Rel. 34, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Calgranulin B (Migration inhibitory factor-related protein 14)
DE   (MRP-14) (p14).
GN   S100A9 OR MRP14.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Lewis/N; TISSUE=Peritoneal cavity;
RX   MEDLINE=93343942; PubMed=8343166;
RA   Imamichi T., Uchida I., Wahl S.M., McCartney-Francis N.;
RT   "Expression and cloning of migration inhibitory factor-related
RT   protein (MRP)8 and MRP14 in arthritis-susceptible rats.";
RL   Biochem. Biophys. Res. Commun. 194:819-825(1993).
RN   [2]
RP   SEQUENCE OF 1-56; 61-64 AND 71-112, ACETYLATION, AND METHYLATION.
RC   TISSUE=Spleen;
RX   MEDLINE=98249881; PubMed=9570842;
RA   Raftery M.J., Geczy C.L.;
RT   "Identification of posttranslational modifications and cDNA sequencing
RT   errors in the rat S100 proteins MRP8 and 14 using electrospray
RT   ionization mass spectrometry.";
RL   Anal. Biochem. 258:285-292(1998).
CC   -!- MASS SPECTROMETRY: MW=13069; MW_ERR=2; METHOD=Electrospray.
CC   -!- SIMILARITY: Belongs to the S-100 family.
CC   -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L18948; AAA18214.1; -.
DR   PIR; JN0686; JN0686.
DR   HSSP; P02638; 1MHO.
DR   InterPro; IPR001751; CaBP_S100.
DR   InterPro; IPR002048; EF-hand.
DR   Pfam; PF00036; efhand; 1.
DR   Pfam; PF01023; S_100; 1.
DR   ProDom; PD003407; CaBP_S100; 1.
DR   PROSITE; PS00018; EF_HAND; FALSE_NEG.
DR   PROSITE; PS00303; S100_CABP; 1.
KW   Calcium-binding; Acetylation; Methylation.
FT   INIT_MET      0      0
FT   MOD_RES       1      1       ACETYLATION.
FT   MOD_RES     106    106       METHYLATION.
FT   CA_BIND      23     36       EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT   CA_BIND      67     78       EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
FT   CONFLICT    105    105       S -> R (IN REF. 1).
SQ   SEQUENCE   112 AA;  13014 MW;  381EEB291175D068 CRC64;

  Query Match             43.8%;  Score 204.5;  DB 1;  Length 112;
  Best Local Similarity   41.8%;  Pred. No. 2.2e-12;
  Matches   38;  Conservative   24;  Mismatches   28;  Indels    1;  Gaps    1;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTK-DQGTIDKIFQN 59
              ::||  :  |||:||||| : || ||| | | |::: |:||| ||  | ::  :  | ::
Db          6 SQLERSISTIINVFHQYSRKYGHPDTLNKAEFKEMVNKDLPNFLKREKRNENLLRDIMED 65

Qy         60 LDANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
              || ||| |:||:| ::|:  ::   |: :|:
Db         66 LDTNQDNQLSFEECMMLMGKLIFACHEKLHE 96
```

```
ESULT 10
10P_HUMAN
D   S10P_HUMAN      STANDARD;      PRT;    95 AA.
C   P25815;
T   01-MAY-1992 (Rel. 22, Created)
T   01-DEC-1992 (Rel. 24, Last sequence update)
T   10-OCT-2003 (Rel. 42, Last annotation update)
E   S-100P protein.
N   S100P OR S100E.
S   Homo sapiens (Human).
C   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
X   NCBI_TaxID=9606;
N   [1]
P   SEQUENCE FROM N.A.
C   TISSUE=Placenta;
X   MEDLINE=92339442; PubMed=1633809;
A   Becker T., Gerke V., Kube E., Weber K.;
T   "S100P, a novel Ca(2+)-binding protein from human placenta. cDNA
T   cloning, recombinant protein expression and Ca2+ binding
T   properties.";
L   Eur. J. Biochem. 207:541-547(1992).
N   [2]
P   SEQUENCE FROM N.A.
C   TISSUE=Placenta;
X   MEDLINE=22388257; PubMed=12477932;
A   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
A   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
A   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
A   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
A   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
A   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
A   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
A   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
A   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
A   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
A   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
A   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
A   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
A   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
A   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
A   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
A   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
T   "Generation and initial analysis of more than 15,000 full-length
T   human and mouse cDNA sequences.";
L   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
N   [3]
P   SEQUENCE OF 1-91.
C   TISSUE=Placenta;
X   MEDLINE=92171935; PubMed=1540168;
A   Emoto Y., Kobayashi R., Akatsuka H., Hidaka H.;
T   "Purification and characterization of a new member of the S-100
T   protein family from human placenta.";
L   Biochem. Biophys. Res. Commun. 182:1246-1253(1992).
C   -!- SUBUNIT: Interacts with S100Z.
C   -!- MISCELLANEOUS: This protein binds two calcium ions.
C   -!- SIMILARITY: Belongs to the S-100 family.
C   -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
C   ---------------------------------------------------------------------
C   This SWISS-PROT entry is copyright. It is produced through a collaboration
C   between the Swiss Institute of Bioinformatics and the EMBL outstation -
C   the European Bioinformatics Institute. There are no restrictions on its
C   use by non-profit institutions as long as its content is in no way
C   modified and this statement is not removed. Usage by and for commercial
C   entities requires a license agreement (See http://www.isb-sib.ch/announce/
C   or send an email to license@isb-sib.ch).
C   ---------------------------------------------------------------------
R   EMBL; X65614; CAA46566.1; -.
R   EMBL; BC006819; AAH06819.1; -.
R   PIR; S24146; S24146.
R   PDB; 1J55; 07-JAN-03.
R   Genew; HGNC:10504; S100P.
DR   MIM; 600614; -.
DR   GO; GO:0005509; F:calcium ion binding; TAS.
DR   GO; GO:0005515; F:protein binding; TAS.
DR   InterPro; IPR001751; CaBP_S100.
DR   InterPro; IPR002048; EF-hand.
DR   Pfam; PF00036; efhand; 1.
DR   Pfam; PF01023; S_100; 1.
DR   ProDom; PD003407; CaBP_S100; 1.
DR   PROSITE; PS00018; EF_HAND; FALSE_NEG.
DR   PROSITE; PS00303; S100_CABP; 1.
KW   Calcium-binding; Placenta; 3D-structure.
FT   CA_BIND      19     32       EF-HAND 1 (LOW AFFINITY).
FT   CA_BIND      62     73       EF-HAND 2 (HIGH AFFINITY).
FT   CONFLICT     32     32       E -> T (IN REF. 3).
FT   CONFLICT     44     44       F -> E (IN REF. 3).
SQ   SEQUENCE   95 AA;  10400 MW;  786E6E3F3EACC6C1 CRC64;

  Query Match             38.8%;  Score 181;  DB 1;  Length 95;
  Best Local Similarity   43.3%;  Pred. No. 2.7e-10;
  Matches   39;  Conservative   15;  Mismatches   36;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              |:||  :  ||::| :||   |   || | ||| |: ||||  |:: ||:  :||: ::|
Db          2 TELETAMGMIIDVFSRYSGSEGSTQTLTKGELKVLMEKELPGFLQSGKDKDAVDKLLKDL 61

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
              ||| | || | ||:| |  :    |    |
Db         62 DANGDAQVDFSEFIVFVAAITSACHKYFEK 91


RESULT 11
S109_MOUSE
ID   S109_MOUSE      STANDARD;      PRT;   112 AA.
AC   P31725;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Calgranulin B (Migration inhibitory factor-related protein 14)
DE   (MRP-14) (P14) (Leukocyte L1 complex heavy chain).
GN   S100A9 OR CAGB OR MRP14.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RX   MEDLINE=92223366; PubMed=1373330;
RA   Lagasse E., Weissman I.L.;
RT   "Mouse MRP8 and MRP14, two intracellular calcium-binding proteins
RT   associated with the development of the myeloid lineage.";
RL   Blood 79:1907-1915(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvJ;
RX   MEDLINE=21102096; PubMed=11169745;
RA   Nacken W.K.F., Lekstrom-Himes J.A., Sorg C., Manitz M.;
RT   "Molecular analysis of the mouse S100A9 gene and evidence that the
RT   myeloid specific transcription factor C/EBPepsilon is not required for
RT   the regulation of the S100A9/A8 gene expression in neutrophils.";
RL   J. Cell. Biochem. 80:606-616(2001).
RN   [3]
RP   SEQUENCE OF 1-9; 75-92 AND 94-108.
RX   MEDLINE=96235204; PubMed=8645219;
RA   Raftery M.J., Harrison C.A., Alewood P., Jones A., Geczy C.L.;
RT   "Isolation of the murine S100 protein MRP14 (14 kDa
RT   migration-inhibitory-factor-related protein) from activated spleen
RT   cells: characterization of post-translational modifications and zinc
RT   binding.";
RL   Biochem. J. 316:285-293(1996).
CC   -!- MASS SPECTROMETRY: MW=12972; MW_ERR=2; METHOD=Electrospray.
CC   -!- MISCELLANEOUS: Has been shown to bind calcium.
CC   -!- SIMILARITY: Belongs to the S-100 family.
```

```
CC   -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M83219; AAB07228.1; -.
DR   EMBL; AJ250496; CAC14292.1; -.
DR   PIR; S68242; S68242.
DR   HSSP; P80511; 1E8A.
DR   MGD; MGI:1338947; S100a9.
DR   InterPro; IPR001751; CaBP_S100.
DR   InterPro; IPR002048; EF-hand.
DR   Pfam; PF00036; efhand; 1.
DR   Pfam; PF01023; S_100; 1.
DR   ProDom; PD003407; CaBP_S100; 1.
DR   ProDom; PD000012; EF-hand; 1.
DR   PROSITE; PS00018; EF_HAND; FALSE_NEG.
DR   PROSITE; PS00303; S100_CABP; 1.
KW   Calcium-binding; Acetylation; Methylation.
FT   INIT_MET      0      0
FT   CA_BIND      23     36       EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT   CA_BIND      67     78       EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
FT   DISULFID     79     90
FT   MOD_RES       1      1       ACETYLATION.
FT   MOD_RES     106    106       METHYLATION.
SQ   SEQUENCE   112 AA;  12918 MW;  A0F810BC175AEC80 CRC64;

  Query Match             38.2%;  Score 178.5;  DB 1;  Length 112;
  Best Local Similarity   36.3%;  Pred. No. 5.6e-10;
  Matches   33;  Conservative   27;  Mismatches   30;  Indels    1;  Gaps    1;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTK-DQGTIDKIFQN 59
              :::|  :  ||: |||||  : || ||| |:| :|::  :|   :|  | ::  |: | ::
Db          6 SQMERSITTIIDTFHQYSRKEGHPDTLSKKEFRQMVEAQLATFMKKEKRNEALINDIMED 65

Qy         60 LDANQDEQVSFKEFVVLVTDVLITAHDNIHK 90
              || ||| |:||:| ::|:  ::   |: :|:
Db         66 LDTNQDNQLSFEECMMLMAKLIFACHEKLHE 96


RESULT 12
S10I_ICTPU
ID   S10I_ICTPU     STANDARD;      PRT;    92 AA.
AC   Q91061;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ictacalcin.
OS   Ictalurus punctatus (Channel catfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
OC   Ictaluridae; Ictalurus.
OX   NCBI_TaxID=7998;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Olfactory epithelium;
RX   MEDLINE=97038343; PubMed=8883937;
RA   Porta A.R., Bettini E., Buiakova O.I., Baker H., Danho W.,
RA   Margolis F.L.;
RT   "Molecular cloning of ictacalcin: a novel calcium-binding protein
RT   from the channel catfish, Ictalurus punctatus.";
RL   Brain Res. Mol. Brain Res. 41:81-89(1996).
CC   -!- FUNCTION: Plays an important role in catfish calcium homeostasis.
CC   -!- TISSUE SPECIFICITY: Abundant in epithelial cells of olfactory
CC       rosette, barbel, skin and gill but not brain or muscle.
CC   -!- SIMILARITY: Belongs to the S-100 family.
CC   -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U33273; AAB52610.1; -.
DR   HSSP; P30801; 1A03.
DR   InterPro; IPR001751; CaBP_S100.
DR   InterPro; IPR002048; EF-hand.
DR   Pfam; PF00036; efhand; 1.
DR   Pfam; PF01023; S_100; 1.
DR   ProDom; PD003407; CaBP_S100; 1.
DR   ProDom; PD000012; EF-hand; 1.
DR   PROSITE; PS00018; EF_HAND; 1.
DR   PROSITE; PS00303; S100_CABP; 1.
KW   Calcium-binding.
FT   CA_BIND      19     32       EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT   CA_BIND      62     73       EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
SQ   SEQUENCE   92 AA;  10021 MW;  DBFE786B7F921C58 CRC64;

  Query Match             35.5%;  Score 166;  DB 1;  Length 92;
  Best Local Similarity   41.5%;  Pred. No. 6.5e-09;
  Matches   34;  Conservative   16;  Mismatches   32;  Indels    0;  Gaps    0;

Qy         10 IINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNLDANQDEQVS 69
              :|: ||:|| : |   || | ||| |:|||      |  || |:||||::|| | |  |
Db         11 LISTFHKYSGKEGDKCTLTKGELKDLLTKELGGAFGNCSDQATLDKIFKDLDTNADGVVD 70

Qy         70 FKEFVVLVTDVLITAHDNIHKE 91
              |:|:  :|    :  : :: |:
Db         71 FQEYATMVACTTMLCNKSLSKK 92


RESULT 13
S10B_RAT
ID   S10B_RAT       STANDARD;      PRT;    91 AA.
AC   P04631;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   S-100 protein, beta chain.
GN   S100B.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85037924; PubMed=6093041;
RA   Kuwano R., Usui H., Maeda T., Fukui T., Yamanari N., Ohtsuka E.,
RA   Ikehara M., Takahashi Y.;
RT   "Molecular cloning and the complete nucleotide sequence of cDNA to
RT   mRNA for S-100 protein of rat brain.";
RL   Nucleic Acids Res. 12:7455-7465(1984).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Kuwano R., Usui H., Maeda T., Araki K., Kurihara T., Yamakuni T.,
RA   Ohtsuka E., Ikehara M., Takahashi Y.;
RT   "Molecular cloning and nucleotide sequences of cDNA and genomic DNA
RT   for alpha and beta subunits of S100 protein.";
RL   Tanaguchi Symp. Brain Sci. 19:243-255(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91359841; PubMed=1653388;
RA   Maeda T., Usui H., Araki K., Kuwano R., Takahashi Y., Suzuki Y.;
RT   "Structure and expression of rat S-100 beta subunit gene.";
RL   Brain Res. Mol. Brain Res. 10:193-202(1991).
RN   [4]
RP   SEQUENCE OF 5-91 FROM N.A.
RX   MEDLINE=87137648; PubMed=3818655;
```

```
RA   Dunn R., Landry C., O'Hanlon D., Dunn J., Allore R., Brown I.,
RA   Marks A.;
RT   "Reduction in S100 protein beta subunit mRNA in C6 rat glioma cells
RT   following treatment with anti-microtubular drugs.";
RL   J. Biol. Chem. 262:3562-3566(1987).
RN   [5]
RP   STRUCTURE BY NMR.
RX   MEDLINE=96387197; PubMed=8794737;
RA   Drohat A.C., Amburgey J.C., Abildgaard F., Starich M.R.,
RA   Baldisseri D.M., Weber D.J.;
RT   "Solution structure of rat apo-S100B(beta beta) as determined by NMR
RT   spectroscopy.";
RL   Biochemistry 35:11577-11588(1996).
RN   [6]
RP   STRUCTURE BY NMR.
RX   MEDLINE=98153156; PubMed=9485423;
RA   Drohat A.C., Baldisseri D.M., Rustandi R.R., Weber D.J.;
RT   "Solution structure of calcium-bound rat S100B(betabeta) as
RT   determined by nuclear magnetic resonance spectroscopy.";
RL   Biochemistry 37:2729-2740(1998).
RN   [7]
RP   STRUCTURE BY NMR.
RX   MEDLINE=99226808; PubMed=10211826;
RA   Drohat A.C., Tjandra N., Baldisseri D.M., Weber D.J.;
RT   "The use of dipolar couplings for determining the solution structure
RT   of rat apo-S100B.";
RL   Protein Sci. 8:800-809(1999).
CC   -!- FUNCTION: Weakly binds calcium but binds zinc very tightly-
CC       distinct binding sites with different affinities exist for both
CC       ions on each monomer. Physiological concentrations of potassium
CC       ion antagonize the binding of both divalent cations, especially
CC       affecting high-affinity calcium-binding sites.
CC   -!- SUBUNIT: Dimer of either two alpha chains, or two beta chains, or
CC       one alpha and one beta chain.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- TISSUE SPECIFICITY: Although predominant among the water-soluble
CC       brain proteins, S-100 is also found in a variety of other tissues.
CC   -!- SIMILARITY: Belongs to the S-100 family.
CC   -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X01090; CAA25567.1; -.
DR   EMBL; M54919; AAA42096.1; -.
DR   EMBL; S53527; -; NOT_ANNOTATED_CDS.
DR   EMBL; S53522; -; NOT_ANNOTATED_CDS.
DR   EMBL; M15705; -; NOT_ANNOTATED_CDS.
DR   PIR; A60046; A26557.
DR   PDB; 1SYM; 07-DEC-96.
DR   PDB; 1QLK; 11-NOV-98.
DR   PDB; 1B4C; 30-DEC-98.
DR   PDB; 1DT7; 26-JUL-00.
DR   PDB; 1MWN; 18-DEC-02.
DR   GO; GO:0005737; C:cytoplasm; ISS.
DR   GO; GO:0005576; C:extracellular; ISS.
DR   GO; GO:0005509; F:calcium ion binding; ISS.
DR   GO; GO:0019210; F:kinase inhibitor activity; ISS.
DR   GO; GO:0048156; F:tau protein binding; ISS.
DR   GO; GO:0008270; F:zinc ion binding; ISS.
DR   GO; GO:0048143; P:astrocyte activation; ISS.
DR   GO; GO:0007409; P:axonogenesis; ISS.
DR   GO; GO:0006874; P:calcium ion homeostasis; ISS.
DR   GO; GO:0006112; P:energy reserve metabolism; ISS.
DR   GO; GO:0048151; P:hyperphosphorylation; ISS.
DR   GO; GO:0006917; P:induction of apoptosis; ISS.
DR   GO; GO:0007611; P:learning and/or memory; ISS.
DR   GO; GO:0045917; P:positive regulation of complement activation; ISS.
DR   GO; GO:0042035; P:regulation of cytokine biosynthesis; ISS.
DR   GO; GO:0048169; P:regulation of long-term neuronal synaptic p. . .; ISS.
DR   GO; GO:0006417; P:regulation of protein biosynthesis; ISS.
DR   InterPro; IPR001751; CaBP_S100.
DR   InterPro; IPR002048; EF-hand.
DR   Pfam; PF00036; efhand; 1.
DR   Pfam; PF01023; S_100; 1.
DR   ProDom; PD003407; CaBP_S100; 1.
DR   PROSITE; PS00018; EF_HAND; 1.
DR   PROSITE; PS00303; S100_CABP; 1.
KW   Calcium-binding; Zinc; Metal-binding; 3D-structure.
FT   INIT_MET      0      0
FT   CA_BIND      18     31       EF-HAND 1 (LOW AFFINITY).
FT   CA_BIND      61     72       EF-HAND 2 (HIGH AFFINITY).
FT   HELIX         2     17
FT   TURN         18     19
FT   STRAND       27     27
FT   HELIX        29     39
FT   TURN         41     42
FT   HELIX        43     46
FT   HELIX        50     61
FT   TURN         62     62
FT   STRAND       68     68
FT   HELIX        70     83
FT   TURN         87     88
SQ   SEQUENCE   91 AA;  10613 MW;  2378AA8B8FF7134D CRC64;

  Query Match             35.1%;  Score 164;  DB 1;  Length 91;
  Best Local Similarity   37.4%;  Pred. No. 9.8e-09;
  Matches   34;  Conservative   19;  Mismatches   38;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ::||  :  :|::|||||| | |    | | |||:||  || : |:  |:|  :||: : |
Db          1 SELEKAMVALIDVFHQYSGREGDKHKLKKSELKELINNELSHFLEEIKEQEVVDKVMETL 60

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              | : | :  |:||:  |: |    |:    |
Db         61 DEDGDGECDFQEFMAFVSMVTTACHEFFEHE 91


RESULT 14
S10Z_HUMAN
ID   S10Z_HUMAN     STANDARD;      PRT;    98 AA.
AC   Q8WXG8;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   S-100Z protein.
GN   S100Z.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., SEQUENCE OF 1-20, SUBUNIT, CALCIUM-BINDING,
RP   TISSUE SPECIFICITY, AND INTERACTION WITH S100P.
RC   TISSUE=Prostate;
RX   MEDLINE=21614385; PubMed=11747429;
RA   Gribenko A.V., Hopper J.E., Makhatadze G.I.;
RT   "Molecular characterization and tissue distribution of a novel member
RT   of the S100 family of EF-hand proteins.";
RL   Biochemistry 40:15538-15548(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
```

```
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SUBUNIT: Homodimer. Interacts with S100P.
CC   -!- TISSUE SPECIFICITY: Highest level of expression in spleen and
CC       leukocytes.
CC   -!- MISCELLANEOUS: This protein binds two calcium ions.
CC   -!- SIMILARITY: Belongs to the S-100 family.
CC   -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF437876; AAL30893.1; -.
DR   EMBL; BC022320; AAH22320.1; -.
DR   InterPro; IPR001751; CaBP_S100.
DR   InterPro; IPR002048; EF-hand.
DR   Pfam; PF00036; efhand; 1.
DR   Pfam; PF01023; S_100; 1.
DR   ProDom; PD003407; CaBP_S100; 1.
DR   ProDom; PD000012; EF-hand; 1.
DR   PROSITE; PS00018; EF_HAND; 1.
DR   PROSITE; PS00303; S100_CABP; 1.
KW   Calcium-binding.
FT   INIT_MET      0      0
FT   CA_BIND      19     32       EF-HAND 1 (LOW AFFINITY).
FT   CA_BIND      62     73       EF-HAND 2 (HIGH AFFINITY).
SQ   SEQUENCE   98 AA;  11430 MW;  A2053E1809E0F401 CRC64;

  Query Match             35.1%;  Score 164;  DB 1;  Length 98;
  Best Local Similarity   41.9%;  Pred. No. 1.1e-08;
  Matches   36;  Conservative   17;  Mismatches   33;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              |:||   :: :| |||:|| :       | | ||| |: :||    |    |:   :||| |:|
Db          2 TQLEMAMDTMIRIFHRYSGKARKRFKLSKGELKLLLQRELTEFLSCQKETQLVDKIVQDL 61

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHD 86
              |||:| :| | ||||:|  : :  :|
Db         62 DANKDNEVDFNEFVVMVAALTVACND 87


RESULT 15
S108_MOUSE
ID   S108_MOUSE     STANDARD;      PRT;    88 AA.
AC   P27005; P31724;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Calgranulin A (Migration inhibitory factor-related protein 8) (MRP-8)
DE   (P8) (Leukocyte L1 complex light chain) (Chemotactic cytokine CP-10)
DE   (Pro-inflammatory S100 cytokine).
GN   S100A8 OR CAGA OR MRP8.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RX   MEDLINE=92223366; PubMed=1373330;
RA   Lagasse E., Weissman I.L.;
RT   "Mouse MRP8 and MRP14, two intracellular calcium-binding proteins
RT   associated with the development of the myeloid lineage.";
RL   Blood 79:1907-1915(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93203618; PubMed=8454868;
RA   Lackmann M., Rajasekariah P., Iismaa S.E., Jones G., Cornish C.J.,
RA   Hu S., Simpson R.J., Moritz R.L., Geczy C.L.;
RT   "Identification of a chemotactic domain of the pro-inflammatory S100
RT   protein CP-10.";
RL   J. Immunol. 150:2981-2991(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvJ; TISSUE=Liver;
RX   MEDLINE=96190772; PubMed=8611640;
RA   Nacken W.K.F., Manitz M.P., Sorg C.;
RT   "Molecular characterisation of the genomic locus of the mouse MRP8
RT   gene.";
RL   Biochim. Biophys. Acta 1315:1-5(1996).
RN   [4]
RP   SEQUENCE OF 1-76.
RC   TISSUE=Spleen;
RX   MEDLINE=92218405; PubMed=1559987;
RA   Lackmann M., Cornish C.J., Simpson R.J., Moritz R.L., Geczy C.L.;
RT   "Purification and structural analysis of a murine chemotactic
RT   cytokine (CP-10) with sequence homology to S100 proteins.";
RL   J. Biol. Chem. 267:7499-7504(1992).
CC   -!- FUNCTION: Has maximal chemotactic activity for neutrophils. May
CC       play an important role in cellular immune reactions.
CC   -!- SUBUNIT: Homodimer and/or heterodimer.
CC   -!- MISCELLANEOUS: Binds two calcium ions per molecule with an
CC       affinity similar to that of the S-100 proteins (By similarity).
CC   -!- SIMILARITY: Belongs to the S-100 family.
CC   -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M83218; AAB07229.1; -.
DR   EMBL; S57123; AAB25840.1; -.
DR   EMBL; X87966; CAA61204.1; -.
DR   PIR; I56163; I56163.
DR   HSSP; P05109; 1MR8.
DR   MGD; MGI:88244; S100a8.
DR   InterPro; IPR001751; CaBP_S100.
DR   InterPro; IPR002048; EF-hand.
DR   Pfam; PF00036; efhand; 1.
DR   Pfam; PF01023; S_100; 1.
DR   ProDom; PD003407; CaBP_S100; 1.
DR   PROSITE; PS00018; EF_HAND; 1.
DR   PROSITE; PS00303; S100_CABP; 1.
KW   Calcium-binding; Chemotaxis.
FT   INIT_MET      0      0
FT   CA_BIND      19     32       EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT   CA_BIND      58     69       EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
FT   CONFLICT     56     56       E -> D (IN REF. 3).
SQ   SEQUENCE   88 AA;  10163 MW;  9B32BDBE53CFE728 CRC64;

  Query Match             34.9%;  Score 163;  DB 1;  Length 88;
  Best Local Similarity   33.0%;  Pred. No. 1.2e-08;
  Matches   30;  Conservative   25;  Mismatches   32;  Indels    4;  Gaps    1;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ::||  |  :|:::| ||   |::  | | : |:::| | |  ::|      |: :|: |
```

```
Db          2 SELEKALSNLIDVYHNYSNIQGNHHALYKNDFKKMVTTECPQFVQNI----NIENLFREL 57

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              | | |  ::|:||: :|  | : :| : |||
Db         58 DINSDNAINFEEFLAMVIKVGVASHKDSHKE 88

Search completed: March  5, 2004, 09:47:10
Job time : 15 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 5, 2004, 09:43:08 ; Search time 39 Seconds
(without alignments)
736.209 Million cell updates/sec

Title: US-09-646-651C-1
Perfect score: 467
Sequence: 1 TKLEDHLEGIINIFHQYSVR.....EFVVLVTDVLITAHDNIHKE 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	289	61.9	70	6	Q9TR16	Q9tr16 bos taurus
2	193	41.3	100	13	Q7ZVA4	Q7zva4 brachydanio
3	178	38.1	99	13	Q8AYJ2	Q8ayj2 squalus aca
4	178	38.1	101	13	O93395	O93395 salvelinus
5	167	35.8	92	11	Q925T3	Q925t3 cricetulus
6	164	35.1	92	13	Q9PSF6	Q9psf6 ictalurus p
7	163	34.9	94	11	Q9D3M4	Q9d3m4 mus musculu
8	163	34.9	94	11	Q91V77	Q91v77 m 11 days e
9	162	34.7	79	11	Q9JL08	Q9jl08 mus musculu
10	154	33.0	148	11	Q8BLX1	Q8blx1 mus musculu
11	154	33.0	2496	11	Q8VHD8	Q8vhd8 mus musculu
12	151	32.3	591	4	Q01720	Q01720 homo sapien
13	151	32.3	687	4	Q9H4U2	Q9h4u2 homo sapien
14	148	31.7	101	6	Q9TV56	Q9tv56 canis famil
15	147	31.5	1218	4	Q05331	Q05331 homo sapien
16	145	31.0	213	4	Q9H4U1	Q9h4u1 homo sapien
17	140.5	30.1	83	11	Q9QVR5	Q9qvr5 rattus sp.
18	139.5	29.9	79	6	Q865V3	Q865v3 equus cabal
19	139	29.8	103	6	Q862H7	Q862h7 bos taurus
20	132	28.3	501	11	Q8CIU0	Q8ciu0 rattus norv
21	129	27.6	357	11	Q8BHC3	Q8bhc3 mus musculu
22	129	27.6	554	11	Q8K552	Q8k552 mus musculu
23	128	27.4	89	11	Q9R2B7	Q9r2b7 rattus norv
24	127.5	27.3	217	13	Q9YH57	Q9yh57 rana catesb
25	126	27.0	80	11	Q91XG5	Q91xg5 mus musculu
26	124	26.6	495	4	Q9UBG3	Q9ubg3 homo sapien
27	124	26.6	495	4	Q8N613	Q8n613 homo sapien
28	112.5	24.1	73	6	Q28714	Q28714 oryctolagus
29	112	24.0	77	6	Q862V4	Q862v4 bos taurus
30	105	22.5	55	6	Q9TSB1	Q9tsb1 bos taurus
31	101.5	21.7	124	11	Q9D708	Q9d708 mus musculu
32	90.5	19.4	170	13	Q7T063	Q7t063 xenopus lae
33	88	18.8	638	11	Q9D3P1	Q9d3p1 mus musculu
34	87.5	18.7	170	5	Q9NKW7	Q9nkw7 patinopecte
35	84.5	18.1	535	10	Q38873	Q38873 arabidopsis
36	82	17.6	110	4	O75544	O75544 homo sapien
37	82	17.6	169	5	Q9NFN1	Q9nfn1 schistosoma
38	81.5	17.5	536	10	O65644	O65644 arabidopsis
39	81	17.3	160	5	Q9BJE9	Q9bje9 plasmodium
40	81	17.3	170	5	Q95P81	Q95p81 bombyx mori
41	80.5	17.2	248	10	O65587	O65587 arabidopsis
42	79.5	17.0	512	10	Q9SE24	Q9se24 oryza sativ
43	78.5	16.8	115	11	Q99LQ9	Q99lq9 mus musculu
44	78.5	16.8	171	5	Q20804	Q20804 caenorhabdi
45	78.5	16.8	523	10	Q9FKW4	Q9fkw4 arabidopsis

ALIGNMENTS

RESULT 1
Q9TR16
ID Q9TR16 PRELIMINARY; PRT; 70 AA.
AC Q9TR16;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CORNEA-associated antigen, CO-AG=CALGRANULIN C homolog.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=96181454; PubMed=8603881;
RA Liu S.H., Gottsch J.D.;
RT "Amino acid sequence of an immunogenic corneal stromal protein.";
RL Invest. Ophthalmol. Vis. Sci. 37:944-948(1996).
CC -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR HSSP; P80511; 1E8A.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF01023; S_100; 1.
DR ProDom; PD003407; CaBP_S100; 1.
SQ SEQUENCE 70 AA; 8134 MW; 7D52BEA97A4D53A5 CRC64;

Query Match 61.9%; Score 289; DB 6; Length 70;
Best Local Similarity 81.4%; Pred. No. 3.3e-20;
Matches 57; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

```
Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              |||||||||||||||||||:||:||| |||||||||||| ||:||||| |||||||:|
Db          1 TKLEDHLEGIINIFHQYSVRVGHFDTLNKRELKQLITKELPKTLQNTKDQPTIDKIFQDL 60

Qy         61 DANQDEQVSF 70
              ||::  | |
Db         61 DADKKGAVVF 70
```

```
RESULT 2
Q7ZVA4
ID   Q7ZVA4      PRELIMINARY;      PRT;   100 AA.
AC   Q7ZVA4;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Body;
RA   Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC045941; AAH45941.1; -.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   InterPro; IPR001751; CaBP_S100.
DR   InterPro; IPR002048; EF-hand.
DR   Pfam; PF00036; efhand; 1.
DR   Pfam; PF01023; S_100; 1.
DR   ProDom; PD003407; CaBP_S100; 1.
DR   SMART; SM00054; EFh; 1.
DR   PROSITE; PS00018; EF_HAND; 1.
DR   PROSITE; PS00303; S100_CABP; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   100 AA;  11157 MW;  DAB81814E54CB8C8 CRC64;

  Query Match             41.3%;  Score 193;  DB 13;  Length 100;
  Best Local Similarity   46.1%;  Pred. No. 6.3e-11;
  Matches   41;  Conservative   17;  Mismatches   31;  Indels    0;  Gaps    0;

Qy          3 LEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNLDA 62
              ||  :| :| :||:||   |:  || :||||||: ||| : ||: ||   :||| ::|||
Db          5 LERAMETLITVFHRYSGAEGNSSTLSRRELKQLMEKELASFLKSQKDPAAVDKIMKDLDA 64

Qy         63 NQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              | | :|:|:||| ||  : |      |:
Db         65 NGDGEVNFEEFVSLVVGLSIACEQLYQKQ 93


RESULT 3
Q8AYJ2
ID   Q8AYJ2      PRELIMINARY;      PRT;    99 AA.
AC   Q8AYJ2;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   S-100 calcium-binding protein A1.
OS   Squalus acanthias (Spiny dogfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Squalea; Squaloidei; Squalidae; Squalus.
OX   NCBI_TaxID=7797;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wang C., Callard G.V.;
RT   "Molecular cloning and stage dependence of an S-10 cDNA from the shark
RT   testis.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF421551; AAN63527.1; -.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   InterPro; IPR001751; CaBP_S100.
DR   InterPro; IPR002048; EF-hand.
DR   Pfam; PF00036; efhand; 1.
DR   Pfam; PF01023; S_100; 1.
DR   ProDom; PD003407; CaBP_S100; 1.
DR   ProDom; PD000012; EF-hand; 1.
DR   SMART; SM00054; EFh; 1.
DR   PROSITE; PS00018; EF_HAND; 1.
DR   PROSITE; PS00303; S100_CABP; 1.
SQ   SEQUENCE   99 AA;  11050 MW;  BA62D8190A4A3693 CRC64;

  Query Match             38.1%;  Score 178;  DB 13;  Length 99;
  Best Local Similarity   40.7%;  Pred. No. 1.7e-09;
  Matches   35;  Conservative   20;  Mismatches   31;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              |:||  : ||| :| :|| : |   :|   |:  |:  |||| ||: ||:  :||| ::|
Db          2 TELESAMAGIIGVFRKYSGKEGDKYSLSNNEMVDLLKAELPNFLKSQKDKAAVDKIMKDL 61

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHD 86
              | |:| :: |:|||||:  :    :|
Db         62 DRNKDGELDFQEFVVLIAALAAACND 87


RESULT 4
O93395
ID   O93395      PRELIMINARY;      PRT;   101 AA.
AC   O93395;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   S100-like calcium binding protein.
GN   S100.
OS   Salvelinus fontinalis (Brook trout) (Brook char).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
OX   NCBI_TaxID=8038;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovulatory ovary;
RX   MEDLINE=20534789; PubMed=11080585;
RA   Bobe J., Goetz F.W.;
RT   "A S100 homologue mRNA isolated by differential display PCR is down-
RT   regulated in the brook trout (Salvelinus fontinalis) post-ovulatory
RT   ovary.";
RL   Gene 257:187-194(2000).
CC   -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR   EMBL; AF077613; AAC28367.1; -.
DR   HSSP; P02633; 4ICB.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   InterPro; IPR001751; CaBP_S100.
DR   InterPro; IPR002048; EF-hand.
DR   Pfam; PF00036; efhand; 1.
DR   Pfam; PF01023; S_100; 1.
DR   ProDom; PD003407; CaBP_S100; 1.
DR   ProDom; PD000012; EF-hand; 1.
DR   PROSITE; PS00018; EF_HAND; 1.
DR   PROSITE; PS00303; S100_CABP; 1.
SQ   SEQUENCE   101 AA;  11285 MW;  BF45582FF9279D0A CRC64;

  Query Match             38.1%;  Score 178;  DB 13;  Length 101;
  Best Local Similarity   45.1%;  Pred. No. 1.7e-09;
  Matches   37;  Conservative   19;  Mismatches   26;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ::||  :| :| :||:|: : |   :|| |:|||:|:  || : ||: ||   || | ::|
Db          3 SQLESSMESLITVFHRYADKDGDCNTLSKKELKELMQTELASFLKSQKDPAAIDTIMKDL 62

Qy         61 DANQDEQVSFKEFVVLVTDVLI 82
              | | | :|||:||| ||  : |
Db         63 DQNGDGKVSFEEFVSLVVGLSI 84


RESULT 5
Q925T3
ID   Q925T3      PRELIMINARY;      PRT;    92 AA.
AC   Q925T3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
```

```
)T   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
)E   S100B.
)S   Cricetulus griseus (Chinese hamster).
)C   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
)C   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
)C   Cricetulus.
)X   NCBI_TaxID=10029;
)N   [1]
)P   SEQUENCE FROM N.A.
)A   Kuge O., Yamakawa Y., Nishijima M.;
)T   "Enhancement of transport-dependent decarboxylation of
)T   phosphatidylserine by S100B protein in permeabilized Chinese hamster
)T   ovary cells.";
)L   J. Biol. Chem. 0:0-0(2001).
)C   -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.
)R   EMBL; AB056121; BAB43945.1; -.
)R   GO; GO:0005737; C:cytoplasm; ISS.
)R   GO; GO:0005576; C:extracellular; ISS.
)R   GO; GO:0005509; F:calcium ion binding; ISS.
)R   GO; GO:0019210; F:kinase inhibitor activity; ISS.
)R   GO; GO:0042803; F:protein homodimerization activity; ISS.
)R   GO; GO:0048155; F:S100 alpha binding; ISS.
)R   GO; GO:0048154; F:S100 beta binding; ISS.
)R   GO; GO:0048156; F:tau protein binding; ISS.
)R   GO; GO:0008270; F:zinc ion binding; ISS.
)R   GO; GO:0048143; P:astrocyte activation; ISS.
)R   GO; GO:0007409; P:axonogenesis; ISS.
)R   GO; GO:0006874; P:calcium ion homeostasis; ISS.
)R   GO; GO:0006112; P:energy reserve metabolism; ISS.
)R   GO; GO:0048151; P:hyperphosphorylation; ISS.
)R   GO; GO:0006917; P:induction of apoptosis; ISS.
)R   GO; GO:0007611; P:learning and/or memory; ISS.
)R   GO; GO:0045917; P:positive regulation of complement activation; ISS.
)R   GO; GO:0042035; P:regulation of cytokine biosynthesis; ISS.
)R   GO; GO:0048169; P:regulation of long-term neuronal synaptic p. . .; ISS.
)R   GO; GO:0006417; P:regulation of protein biosynthesis; ISS.
)R   InterPro; IPR001751; CaBP_S100.
)R   InterPro; IPR002048; EF-hand.
)R   Pfam; PF00036; efhand; 1.
)R   Pfam; PF01023; S_100; 1.
)R   ProDom; PD003407; CaBP_S100; 1.
)R   PROSITE; PS00018; EF_HAND; 1.
)R   PROSITE; PS00303; S100_CABP; 1.
)Q   SEQUENCE   92 AA;  10749 MW;  AF50107EC2BEDF6B CRC64;

 Query Match             35.8%;  Score 167;  DB 11;  Length 92;
 Best Local Similarity   38.5%;  Pred. No. 1.7e-08;
 Matches   35;  Conservative   19;  Mismatches   37;  Indels    0;  Gaps    0;

)y          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ::||  :  :|:|||||| | |     | | |||:||  || : |:  |:|  :||: : |
)b        .     2 SELEKAMVALIDIFHQYSGREGDKHKLKKSELKELINNELSHFLEEIKEQEVVDKVMETL 61

)y         61 DANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              | : | :  |:||:  |: |    |:   :|
)b         62 DEDGDGECDFQEFMAFVSMVTTACHEFFEQE 92


RESULT 6
Q9PSF6
ID   Q9PSF6      PRELIMINARY;      PRT;    92 AA.
AC   Q9PSF6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Ictacalcin.
OS   Ictalurus punctatus (Channel catfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
OC   Ictaluridae; Ictalurus.
OX   NCBI_TaxID=7998;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94376615; PubMed=8090068;
RA   Bettini E., Porta A.R., Dahmen N., Wang H., Margolis F.L.;
RT   "Expressed sequence tags (EST) identify genes preferentially expressed
RT   in catfish chemosensory tissues.";
RL   Brain Res. Mol. Brain Res. 23:285-291(1994).
CC   -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR   HSSP; P30801; 1A03.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   InterPro; IPR001751; CaBP_S100.
DR   InterPro; IPR002048; EF-hand.
DR   Pfam; PF00036; efhand; 1.
DR   Pfam; PF01023; S_100; 1.
DR   ProDom; PD003407; CaBP_S100; 1.
DR   ProDom; PD000012; EF-hand; 1.
DR   PROSITE; PS00018; EF_HAND; 1.
DR   PROSITE; PS00303; S100_CABP; 1.
SQ   SEQUENCE   92 AA;  10022 MW;  E04875D0C9921C50 CRC64;

  Query Match             35.1%;  Score 164;  DB 13;  Length 92;
  Best Local Similarity   41.5%;  Pred. No. 3.2e-08;
  Matches   34;  Conservative   15;  Mismatches   33;  Indels    0;  Gaps    0;

Qy          10 IINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNLDANQDEQVS 69
               :|: ||:||    |   || | ||| |:||||      |  || |:||||::|| | |  |
Db          11 LISTFHKYSGEEGDKCTLTKGELKDLLTKELGGAFGNCSDQATLDKIFKDLDTNADGVVD 70

Qy          70 FKEFVVLVTDVLITAHDNIHKE 91
               |:|:  :|    :  : :: |:
Db          71 FQEYATMVACTTMLCNKSLSKK 92


RESULT 7
Q9D3M4
ID   Q9D3M4      PRELIMINARY;      PRT;    94 AA.
AC   Q9D3M4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   S100 calcium binding protein A1.
GN   S100A1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Head;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
CC   -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR   EMBL; AK017279; BAB30670.1; -.
DR   HSSP; P04631; 1B4C.
DR   MGD; MGI:1338917; S100a1.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
```

```
DR   InterPro; IPR001751; CaBP_S100.
DR   InterPro; IPR002048; EF-hand.
DR   Pfam; PF00036; efhand; 1.
DR   Pfam; PF01023; S_100; 1.
DR   ProDom; PD003407; CaBP_S100; 1.
DR   ProDom; PD000012; EF-hand; 1.
DR   PROSITE; PS00018; EF_HAND; 1.
DR   PROSITE; PS00303; S100_CABP; 1.
SQ   SEQUENCE   94 AA;  10487 MW;  7C8E817D135C2D15 CRC64;

  Query Match             34.9%;  Score 163;  DB 11;  Length 94;
  Best Local Similarity   39.5%;  Pred. No. 4.1e-08;
  Matches   34;  Conservative   18;  Mismatches   34;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ::||  :| :||:|| :| : |    | |:||| |:  ||   |   ||   :||: : |
Db          3 SELESAMETLINVFHAHSGKEGDKYKLSKKELKDLLQTELSGFLDVQKDADAVDKVIKEL 62

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHD 86
              | | | :| |||:||||  : :  ::
Db         63 DENGDGEVDFKEYVVLVAALTVACNN 88


RESULT 8
Q91V77
ID   Q91V77      PRELIMINARY;      PRT;    94 AA.
AC   Q91V77;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   11 days embryo cDNA, RIKEN full-length enriched library,
DE   clone:2700088D09, full insert sequence (S100 calcium binding protein
DE   A1) (Adult male kidney cDNA, RIKEN full-length enriched library,
DE   clone:0610031F03, full insert sequence).
GN   S100A1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Kidney;
RA   Adachi J., Aizawa K., Akahira S., Akimura T., Aono H., Arai A.,
RA   Arakawa T., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA   Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA   Imotani K., Ishii Y., Itoh M., Izawa M., Kato H., Kawai J., Kojima Y.,
RA   Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A.,
RA   Nishi K., Nomura K., Numazaki R., Ohno M., Okazaki Y., Okido T.,
RA   Owa C., Saito H., Saito R., Sakai C., Sakai K., Sano H., Sasaki D.,
RA   Shibata K., Shibata Y., Shinagawa A., Shiraki T., Sogabe Y.,
RA   Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T., Tejima Y.,
RA   Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA   Muramatsu M., Hayashizaki Y.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Kidney;
RX   MEDLINE=21085660; PubMed=11217851;
RA   RIKEN FANTOM Consortium.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Kidney;
RX   MEDLINE=99279253; PubMed=10349636;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Kidney;
RX   MEDLINE=20499374; PubMed=11042159;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Kidney;
RX   MEDLINE=20530913; PubMed=11076861;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/Sv;
RA   Du X.-J., Cole T.J., Tenis N., Gao X.-M., Kontgen F., Kemp B.E.,
RA   Heierhorst J.;
RT   "Impaired cardiac contractility without cardiomyopathy in S100A1-
RT   deficient mice.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR   EMBL; AK012578; BAB28330.1; -.
DR   EMBL; BC005590; AAH05590.1; -.
DR   EMBL; AF368423; AAL14436.1; -.
DR   EMBL; AK002721; BAB22308.1; -.
DR   MGD; MGI:1338917; S100a1.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   InterPro; IPR001751; CaBP_S100.
DR   InterPro; IPR002048; EF-hand.
DR   Pfam; PF00036; efhand; 1.
DR   Pfam; PF01023; S_100; 1.
DR   ProDom; PD003407; CaBP_S100; 1.
DR   ProDom; PD000012; EF-hand; 1.
DR   PROSITE; PS00018; EF_HAND; 1.
DR   PROSITE; PS00303; S100_CABP; 1.
SQ   SEQUENCE   94 AA;  10505 MW;  7C8E817BBF36ED15 CRC64;

  Query Match             34.9%;  Score 163;  DB 11;  Length 94;
  Best Local Similarity   39.5%;  Pred. No. 4.1e-08;
  Matches   34;  Conservative   18;  Mismatches   34;  Indels    0;  Gaps    0;

Qy          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
              ::||  :| :||:|| :| : |    | |:||| |:  ||   |   ||   :||: : |
Db          3 SELESAMETLINVFHAHSGKEGDKYKLSKKELKDLLQTELSGFLDVQKDADAVDKVMKEL 62

Qy         61 DANQDEQVSFKEFVVLVTDVLITAHD 86
              | | | :| |||:||||  : :  ::
Db         63 DENGDGEVDFKEYVVLVAALTVACNN 88


RESULT 9
Q9JL08
ID   Q9JL08      PRELIMINARY;      PRT;    79 AA.
AC   Q9JL08;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   S100 calcium binding protein A1 (Fragment).
GN   S100A1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
```

```
N   [1]
P   SEQUENCE FROM N.A.
A   Kiewitz R., Lyons G.E., Schafer B.W., Heizmann C.W.;
T   "Transcriptional regulation of S100A1 and expression during mouse
T   heart development.";
L   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
C   -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.
R   EMBL; AF218353; AAF32320.1; -.
R   HSSP; P04631; 1B4C.
R   MGD; MGI:1338917; S100a1.
R   GO; GO:0005509; F:calcium ion binding; IEA.
R   InterPro; IPR001751; CaBP_S100.
R   InterPro; IPR002048; EF-hand.
R   Pfam; PF00036; efhand; 1.
R   Pfam; PF01023; S_100; 1.
R   ProDom; PD003407; CaBP_S100; 1.
R   ProDom; PD000012; EF-hand; 1.
R   PROSITE; PS00018; EF_HAND; 1.
R   PROSITE; PS00303; S100_CABP; 1.
T   NON_TER      79     79
Q   SEQUENCE   79 AA;  8863 MW;  F94EDA3A798815D5 CRC64;

 Query Match             34.7%;  Score 162;  DB 11;  Length 79;
 Best Local Similarity   44.2%;  Pred. No. 4.3e-08;
 Matches   34;  Conservative   14;  Mismatches   29;  Indels    0;  Gaps    0;

y          1 TKLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
             ::||  :| :||:|| :| : |    | |:||| |:  ||   |   ||   :||: : |
b          3 SELESAMETLINVFHAHSAQEGDKYKLSKKELKDLLQTELSGFLDVQKDADAVDKVMKEL 62

y         61 DANQDEQVSFKEFVVLV 77
             | | | :| |||:||||
b         63 DENGDGEVDFKEYVVLV 79


ESULT 10
8BLX1
D   Q8BLX1      PRELIMINARY;      PRT;   148 AA.
C   Q8BLX1;
T   01-MAR-2003 (TrEMBLrel. 23, Created)
T   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
T   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
E   Weakly similar to DJ14N1.2 (Fragment).
S   Mus musculus (Mouse).
C   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
X   NCBI_TaxID=10090;
N   [1]
P   SEQUENCE FROM N.A.
C   STRAIN=C57BL/6J; TISSUE=Aorta, and Vein;
X   MEDLINE=22354683; PubMed=12466851;
A   The FANTOM Consortium,
A   the RIKEN Genome Exploration Research Group Phase I & II Team;
T   "Analysis of the mouse transcriptome based on functional annotation of
T   60,770 full-length cDNAs.";
L   Nature 420:563-573(2002).
R   EMBL; AK041026; BAC30787.1; -.
R   GO; GO:0005509; F:calcium ion binding; IEA.
R   InterPro; IPR001751; CaBP_S100.
R   InterPro; IPR002048; EF-hand.
R   Pfam; PF00036; efhand; 1.
R   Pfam; PF01023; S_100; 1.
R   ProDom; PD003407; CaBP_S100; 1.
R   ProDom; PD000012; EF-hand; 1.
R   PROSITE; PS00018; EF_HAND; 1.
R   PROSITE; PS00303; S100_CABP; 1.
T   NON_TER     148    148
Q   SEQUENCE   148 AA;  16934 MW;  590220BA68351058 CRC64;

 Query Match             33.0%;  Score 154;  DB 11;  Length 148;
 Best Local Similarity   33.3%;  Pred. No. 4.8e-07;
 Matches   30;  Conservative   23;  Mismatches   37;  Indels    0;  Gaps    0;

Qy         2 KLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNLD 61
             || : :  :|::|:||:   |: | | | |:|:|:  |    |||  |  |:| | ||||
Db         3 KLLESIVTVIDVFYQYATEYGNCDMLSKEEMKELLVTEFHQILKNPDDPDTVDIIMQNLD 62

Qy        62 ANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
               : : :| | |:::::  :    :  | |:
Db        63 RDHNHKVDFTEYLLMILKLTKACNKIIGKD 92


RESULT 11
Q8VHD8
ID   Q8VHD8      PRELIMINARY;      PRT;  2496 AA.
AC   Q8VHD8;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hornerin.
GN   HORNERIN OR 1110033K19RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ICR;
RX   MEDLINE=21601613; PubMed=11572870;
RA   Makino T., Takaishi M., Morohashi M., Huh N.-H.;
RT   "Hornerin, a Novel Profilaggrin-like Protein and Differentiation-
RT   specific Marker Isolated from Mouse Skin.";
RL   J. Biol. Chem. 276:47445-47452(2001).
CC   -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR   EMBL; AY027660; AAK15791.1; -.
DR   MGD; MGI:1915973; Hornerin.
DR   GO; GO:0001533; C:cornified envelope; IDA.
DR   InterPro; IPR001751; CaBP_S100.
DR   InterPro; IPR002048; EF-hand.
DR   Pfam; PF00036; efhand; 1.
DR   Pfam; PF01023; S_100; 1.
DR   ProDom; PD003407; CaBP_S100; 1.
DR   PROSITE; PS00018; EF_HAND; 1.
DR   PROSITE; PS00303; S100_CABP; 1.
SQ   SEQUENCE   2496 AA;  287588 MW;  4CE136CA6CE657DE CRC64;

  Query Match             33.0%;  Score 154;  DB 11;  Length 2496;
  Best Local Similarity   33.3%;  Pred. No. 1.1e-05;
  Matches   30;  Conservative   23;  Mismatches   37;  Indels    0;  Gaps    0;

Qy         2 KLEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNLD 61
             || : :  :|::|:||:   |: | | | |:|:|:  |    |||  |  |:| | ||||
Db         3 KLLESIVTVIDVFYQYATEYGNCDMLSKEEMKELLVTEFHQILKNPDDPDTVDIIMQNLD 62

Qy        62 ANQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
               : : :| | |:::::  :    :  | |:
Db        63 RDHNHKVDFTEYLLMILKLTKACNKIIGKD 92


RESULT 12
Q01720
ID   Q01720      PRELIMINARY;      PRT;   591 AA.
AC   Q01720;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   FILAGGRIN precursor (PROFILAGGRIN) (Fragment).
GN   FLG.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RX   MEDLINE=93054736; PubMed=1429717;
```

```
RA   Presland R.B., Haydock P.V., Fleckman P., Nirunsuksiri W., Dale B.A.;
RT   "Characterization of the human epidermal profilaggrin gene. Genomic
RT   organization and identification of an S-100-like calcium binding
RT   domain at the amino terminus.";
RL   J. Biol. Chem. 267:23772-23781(1992).
CC   -!- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
CC       DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
CC       TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC   -!- PTM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
CC       HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
CC       324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES.
CC       DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
CC       PROTEOLYTICALLY CLEAVED.
CC   -!- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
CC       WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
CC       REPEATS.
DR   EMBL; L01089; AAA60177.1; -.
DR   EMBL; L01090; AAA60176.1; -.
DR   PIR; A45135; A45135.
DR   PIR; A48118; A48118.
DR   HSSP; P80511; 1E8A.
DR   MIM; 135940; -.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR001751; CaBP_S100.
DR   InterPro; IPR002048; EF-hand.
DR   InterPro; IPR003303; Filaggrin.
DR   Pfam; PF00036; efhand; 1.
DR   Pfam; PF03516; Filaggrin; 2.
DR   Pfam; PF01023; S_100; 1.
DR   PRINTS; PR00487; FILAGGRIN.
DR   PROSITE; PS00018; EF_HAND; 1.
DR   PROSITE; PS00303; S100_CABP; 1.
KW   Polymorphism.
FT   PROPEP          1    293       POTENTIAL.
FT   CHAIN         294    467       FILAGGRIN.
FT   PROPEP        468    474       POTENTIAL.
FT   CHAIN         475   >591       FILAGGRIN.
FT   CA_BIND        19     32       SITE I (BY SIMILARITY).
FT   CA_BIND        62     73       SITE II (BY SIMILARITY).
FT   NON_TER       591    591
SQ   SEQUENCE   591 AA;  66366 MW;  381491625C75E369 CRC64;

  Query Match             32.3%;  Score 151;  DB 4;  Length 591;
  Best Local Similarity   34.8%;  Pred. No. 4.2e-06;
  Matches   31;  Conservative   23;  Mismatches   35;  Indels    0;  Gaps    0;

Qy          3 LEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNLDA 62
              | :::  |||:| |||  :  : ||| |:|||:|: ||    |||  |   :|    :||
Db          4 LLENIFAIINLFKQYSKKDKNTDTLSKKELKELLEKEFRQILKNPDDPDMVDVFMDHLDI 63

Qy         63 NQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              : :::: | ||:::|  :    :::  ||
Db         64 DHNKKIDFTEFLLMVFKLAQAYYESTRKE 92


RESULT 13
Q9H4U2
ID   Q9H4U2      PRELIMINARY;      PRT;   687 AA.
AC   Q9H4U2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   DJ14N1.1.1 (Profilaggrin 5' end) (Fragment).
GN   FLG.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Laird G.;
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR   EMBL; AL356504; CAC13172.1; -.
DR   PIR; A48118; A48118.
DR   HSSP; P80511; 1E8A.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR001751; CaBP_S100.
DR   InterPro; IPR002048; EF-hand.
DR   InterPro; IPR003303; Filaggrin.
DR   Pfam; PF00036; efhand; 1.
DR   Pfam; PF03516; Filaggrin; 3.
DR   Pfam; PF01023; S_100; 1.
DR   PRINTS; PR00487; FILAGGRIN.
DR   PROSITE; PS00018; EF_HAND; 1.
DR   PROSITE; PS00303; S100_CABP; 1.
FT   NON_TER       687    687
SQ   SEQUENCE   687 AA;  76659 MW;  8000363FBEF07B74 CRC64;

  Query Match             32.3%;  Score 151;  DB 4;  Length 687;
  Best Local Similarity   34.8%;  Pred. No. 5e-06;
  Matches   31;  Conservative   23;  Mismatches   35;  Indels    0;  Gaps    0;

Qy          3 LEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNLDA 62
              | :::  |||:| |||  :  : ||| |:|||:|: ||    |||  |   :|    :||
Db          4 LLENIFAIINLFKQYSKKDKNTDTLSKKELKELLEKEFRQILKNPDDPDMVDVFMDHLDI 63

Qy         63 NQDEQVSFKEFVVLVTDVLITAHDNIHKE 91
              : :::: | ||:::|  :    :::  ||
Db         64 DHNKKIDFTEFLLMVFKLAQAYYESTRKE 92


RESULT 14
Q9TV56
ID   Q9TV56      PRELIMINARY;      PRT;   101 AA.
AC   Q9TV56;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Metastasin.
GN   MTS1.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Madin-Durby; TISSUE=Kidney;
RA   Miyamori H., Hasegawa K., Kim K., Sato H.;
RT   "Expression of metastasis associated mts1 gene is co-induced with
RT   membrane type-1 matrix metalloproteinase (MT1-MMP) during oncogenic
RT   transformation and tubular formation of madin darby canine kidney
RT   (MDCK) epithelial cells.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR   EMBL; AB031064; BAA83419.1; -.
DR   HSSP; P30801; 1A03.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   InterPro; IPR001751; CaBP_S100.
DR   InterPro; IPR002048; EF-hand.
DR   Pfam; PF00036; efhand; 1.
DR   Pfam; PF01023; S_100; 1.
DR   ProDom; PD003407; CaBP_S100; 1.
DR   ProDom; PD000012; EF-hand; 1.
DR   PROSITE; PS00018; EF_HAND; 1.
DR   PROSITE; PS00303; S100_CABP; 1.
SQ   SEQUENCE   101 AA;  11847 MW;  5535387DB7577DF0 CRC64;

  Query Match             31.7%;  Score 148;  DB 6;  Length 101;
  Best Local Similarity   34.5%;  Pred. No. 1.2e-06;
  Matches   29;  Conservative   23;  Mismatches   32;  Indels    0;  Gaps    0;

Qy          3 LEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNLDA 62
              ||  |: ::: ||:|| : |    | : |||:|: :|||: |    |:    |: |||:
```

```
b          5 LEKALDVMVSTFHKYSGKEGDKFKLNRSELKELLMRELPSFLGKRTDEAAFQKLMSNLDS 64

y         63 NQDEQVSFKEFVVLVTDVLITAHD 86
             |:| :| |:|: | :: | :  ::
b         65 NRDNEVDFQEYCVFLSCVAMMCNE 88


ESULT 15
05331
D   Q05331      PRELIMINARY;      PRT;  1218 AA.
AC   Q05331;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   FILAGGRIN (PROFILAGGRIN) (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=FORESKIN;
RX   MEDLINE=93109348; PubMed=8417356;
RA   Markova N.G., Marekov L.N., Chipev C.C., Gan S.-Q., Idler W.W.,
RA   Steinert P.M.;
RT   "Profilaggrin is a major epidermal calcium-binding protein.";
RL   Mol. Cell. Biol. 13:613-625(1993).
CC   -!- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
CC       DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
CC       TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC   -!- PTM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
CC       HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
CC       324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES.
CC       DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
CC       PROTEOLYTICALLY CLEAVED.
CC   -!- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
CC       WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
CC       REPEATS.
DR   EMBL; M96943; AAA36487.1; -.
DR   PIR; A48118; A48118.
DR   HSSP; P02593; 1CDM.
DR   GO; GO:0005856; C:cytoskeleton; NAS.
DR   GO; GO:0005509; F:calcium ion binding; TAS.
DR   GO; GO:0030154; P:cell differentiation; NAS.
DR   GO; GO:0008151; P:cell growth and/or maintenance; NAS.
DR   InterPro; IPR001751; CaBP_S100.
DR   InterPro; IPR002048; EF-hand.
DR   InterPro; IPR003303; Filaggrin.
DR   Pfam; PF00036; efhand; 1.
DR   Pfam; PF03516; Filaggrin; 6.
DR   Pfam; PF01023; S_100; 1.
DR   PRINTS; PR00487; FILAGGRIN.
DR   PROSITE; PS00018; EF_HAND; 1.
DR   PROSITE; PS00303; S100_CABP; 1.
KW   Phosphorylation; Polyprotein; Developmental protein; Calcium-binding;
KW   Polymorphism.
FT   CA_BIND      19     32       SITE I (BY SIMILARITY).
FT   CA_BIND      62     73       SITE II (BY SIMILARITY).
FT   NON_TER    1218   1218
SQ   SEQUENCE   1218 AA;  133604 MW;  EC195AD5285B19C2 CRC64;

 Query Match             31.5%;  Score 147;  DB 4;  Length 1218;
 Best Local Similarity   35.4%;  Pred. No. 2.2e-05;
 Matches   29;  Conservative   22;  Mismatches   31;  Indels    0;  Gaps    0;

Qy         10 IINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNLDANQDEQVS 69
              |||:|::|| :  : ||| |:|||:|: ||    |||  |   :|    :|| : ::::
Db         11 IINLFNEYSKKDKNTDTLSKKELKELLEKEFRQILKNPDDPDMVDVFMDHLDIDHNKKID 70

Qy         70 FKEFVVLVTDVLITAHDNIHKE 91
              | ||:::|  :    :::  ||
Db         71 FTEFLLMVFKLAQAYYESTRKE 92
```

Search completed: March 5, 2004, 09:48:53
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

M nucleic - nucleic search, using sw model

un on: March 8, 2004, 03:14:22 ; Search time 2062 Seconds
(without alignments)
2249.129 Million cell updates/sec

itle: US-09-646-651C-2
erfect score: 107
equence: 1 ggaaaaunnnnnunauaugn.....nnnnnnnnnnnnnnuuagcag 107

coring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

earched: 3470272 seqs, 21671516995 residues

otal number of hits satisfying chosen parameters: 6940544

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		31.4	29.3	298	6	A94582	A94582 Sequence 3
2	c	30	28.0	162986	9	CNS05TBQ	AL163953 Human chr
3	c	30	28.0	167398	9	CNS07EET	AL445886 Human chr
4		29.8	27.9	1311	3	BT001281	BT001281 Drosophil
5	c	29.8	27.9	133636	9	AC117436	AC117436 Homo sapi
6		29.6	27.7	86050	9	AL359647	AL359647 Human DNA
7		29.4	27.5	17492	1	U67539	U67539 Methanococc
8	c	29.4	27.5	110000	6	AR271569_05	Continuation (6 of
9	c	29.2	27.3	60471	2	AC015173	AC015173 Drosophil
10	c	29.2	27.3	164038	3	AC009846	AC009846 Drosophil
11	c	29.2	27.3	234378	3	AE003740	AE003740 Drosophil
12		29	27.1	191942	9	AC079951	AC079951 Homo sapi
13	c	28.6	26.7	25233	3	CEC01B9	Z93373 Caenorhabdi
14	c	28.6	26.7	113880	3	PFMAL3P4	AL008970 Plasmodiu
15	c	28.6	26.7	227182	2	AC099283	AC099283 Rattus no
16	c	28.6	26.7	252129	2	AC133975	AC133975 Rattus no
17		28.4	26.5	1818	1	AF267206	AF267206 Candidatu
18		28.4	26.5	112537	9	AC112251	AC112251 Homo sapi
19	c	28.4	26.5	167560	7	AY303349	AY303349 Enterobac
20		28.4	26.5	177897	2	AC021473	AC021473 Homo sapi
21		28.4	26.5	178204	9	AC097467	AC097467 Homo sapi
22		28.2	26.4	49843	3	AF288092	AF288092 Naegleria
23	c	28	26.2	146500	9	AC004453	AC004453 Homo sapi
24	c	28	26.2	216941	2	AC146064	AC146064 Pan trogl
25		28	26.2	263150	2	AC105564	AC105564 Rattus no
26		27.8	26.0	150641	3	CEY41E3	Z95559 Caenorhabdi
27		27.8	26.0	178607	9	AC007068	AC007068 Homo sapi
28		27.8	26.0	181542	2	AC121076	AC121076 Canis fam
29		27.6	25.8	102704	5	AL672216	AL672216 Zebrafish
30	c	27.6	25.8	120039	9	AL357124	AL357124 Human DNA
31	c	27.6	25.8	132254	3	AC116330	AC116330 Dictyoste
32		27.6	25.8	147341	8	AC122143	AC122143 Oryza sat
33	c	27.6	25.8	160404	9	AL159152	AL159152 Human DNA
34		27.6	25.8	166315	9	AC078940	AC078940 Homo sapi
35		27.4	25.6	524	6	AX541033	AX541033 Sequence
36		27.4	25.6	769	6	AX541172	AX541172 Sequence
37		27.4	25.6	789	6	AX541169	AX541169 Sequence
38		27.4	25.6	835	6	AX541247	AX541247 Sequence
39		27.4	25.6	110000	2	PFMAL13_11	Continuation (12 o
40		27.4	25.6	138027	9	AL590227	AL590227 Human DNA
41	c	27.4	25.6	152313	9	AL591594	AL591594 Human DNA
42	c	27.4	25.6	173342	9	AC023796	AC023796 Homo sapi
43	c	27.4	25.6	183602	2	AC024025	AC024025 Homo sapi
44	c	27.4	25.6	195583	9	AC092867	AC092867 Homo sapi
45		27.4	25.6	283919	2	AC095169	AC095169 Rattus no

ALIGNMENTS

RESULT 1
A94582
LOCUS A94582 298 bp DNA linear PAT 26-JAN-2000
DEFINITION Sequence 3 from Patent WO9947561.
ACCESSION A94582
VERSION A94582.1 GI:6778899
KEYWORDS .
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 298)
AUTHORS Koch-Pelster,B. and Kuhn,E.
TITLE RIBONUCLEOTIDE POLYPEPTIDES CONTAINING METAL
JOURNAL Patent: WO 9947561-A 3 23-SEP-1999;
KOCH PELSTER BRIGITTE (DE); KUHN ECKEHARD (DE)

```
FEATURES             Location/Qualifiers
     source          1. .298
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             29.3%;  Score 31.4;  DB 6;  Length 298;
  Best Local Similarity   30.3%;  Pred. No. 3.7e+02;
  Matches   23;  Conservative    9;  Mismatches   44;  Indels    0;  Gaps    0;

Qy         32 UUNNNNNNAAAAANUANAAACAUNNNNNCUUNAGNNNNNNNNNNNNNAGAAAUNNNNNNN 91
              ::       |||| :| |||||:     |:: ||             |||||:
Db         89 TTTGTACATAAAAATAAAAACATGACTTCTTTAGACACTCCTTCATTAGAAATAAAATAA 148

Qy         92 NNNNNNNNNUUAGCAG 107
                        ::|||||
Db        149 AATAAACTATTAGCAG 164


RESULT 2
CNS05TBQ/c
LOCUS       CNS05TBQ              162986 bp    DNA     linear   PRI 03-MAY-2001
DEFINITION  Human chromosome 14 DNA sequence BAC R-533L7 of library RPCI-11
            from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION   AL163953
VERSION     AL163953.3  GI:12001746
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 162986)
  AUTHORS   Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
            Brottier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,
            Levy,M., Eckenberg,R., Bruls,T., deBerardinis,V., Cruaud,C.,
            Gyapay,G., Saurin,W. and Weissenbach,J.
  TITLE     Sequencing of the human chromosome 14
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 162986)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAY-2001) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     On Jan 1, 2001 this sequence version replaced gi:7768128.
            -------------- Genome Center
            Center: Genoscope / Centre National de Sequencage
            Center code: GS
            Web site: http://www.genoscope.cns.fr/
            Contact: SeqRef@genoscope.cns.fr
            --------------
            The following BAC sequence is oriented from the T7 to the SP6 end.
            -------------- Summary Statistics
            Assembly program: Phrap; version 2.0
            Quality coverage: 8.81x in Q20 bases; sum-of-contigs
            --------------
            ----------------------
            Overall quality chart :
            Range     :   bases
            0         :
            1  -  9   :
            10 - 19   :
            20 - 29   :
            30 - 39   :       2
            40 - 49   :     251
            50 - 59   :    1196
            60 - 69   :    2599
            70 - 79   :    9911
            80 - 89   :   40597
            90 - 99   :  108430
            ----------------------
            Percentage of bases with a quality value >= 40 : 99 %.

FEATURES             Location/Qualifiers
     source          1. .162986
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /clone="R-533L7"
                     /clone_lib="RPCI-11"
     STS             133879. .134005
                     /note="matching EMBL:N74379
                     RHdb:RH91341
                     dbSTS:STS64476
                     Identified using the e-PCR software (G. Schuler)"
     STS             151705. .151879
                     /note="matching EMBL:G03450
                     RHdb:RH53483
                     RHdb:RH3638
                     dbSTS:STS817
                     Identified using the e-PCR software (G. Schuler)"
ORIGIN

  Query Match             28.0%;  Score 30;  DB 9;  Length 162986;
  Best Local Similarity   22.1%;  Pred. No. 1.3e+02;
  Matches   23;  Conservative   13;  Mismatches   68;  Indels    0;  Gaps    0;

Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              ||||      :  :|:|      |:   :::      | |   :| | | |     |::
Db     159690 AAAAAAATCATATTATGTATAATCTTTTTTTTTGGTGACAGGCTAGAGAAAAGGATGCTT 159631

Qy         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUAGCA 106
                |               |||||:               ::| ||
Db     159630 GGGTGGCAATTGAAATAGAAATCCACTTTCAAATACTATTACCA 159587


RESULT 3
CNS07EET/c
LOCUS       CNS07EET              167398 bp    DNA     linear   PRI 29-APR-2001
DEFINITION  Human chromosome 14 DNA sequence BAC R-369B6 of library RPCI-11
            from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION   AL445886
VERSION     AL445886.3  GI:13897491
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 167398)
  AUTHORS   Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
            Brottier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,
            Levy,M., Eckenberg,R., Bruls,T., deBerardinis,V., Cruaud,C.,
            Gyapay,G., Saurin,W. and Weissenbach,J.
  TITLE     Sequencing of the human chromosome 14
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 167398)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-APR-2001) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     On Apr 30, 2001 this sequence version replaced gi:11181947.
            -------------- Genome Center
            Center: Genoscope / Centre National de Sequencage
            Center code: GS
            Web site: http://www.genoscope.cns.fr/
            Contact: SeqRef@genoscope.cns.fr
            --------------
            The following BAC sequence is oriented from the T7 to the SP6 end.
            -------------- Summary Statistics
            Assembly program: Phrap; version 2.0
            Quality coverage: 7.31x in Q20 bases; sum-of-contigs
            --------------
            ----------------------
            Overall quality chart :
```

```
           Range    :   bases
           0        :
           1  -  9  :
           10 - 19  :      62
           20 - 29  :     234
           30 - 39  :     702
           40 - 49  :    3036
           50 - 59  :    5255
           60 - 69  :    8791
           70 - 79  :   18773
           80 - 89  :   41646
           90 - 99  :   88899
           ------------------------
           Percentage of bases with a quality value >= 40 : 99 %.
'EATURES             Location/Qualifiers
     source          1. .167398
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /clone="R-369B6"
                     /clone_lib="RPCI-11"
     STS             90076. .90202
                     /note="matching EMBL:N74379
                     RHdb:RH91341
                     dbSTS:STS64476
                     Identified using the e-PCR software (G. Schuler)"
     STS             107900. .108074
                     /note="matching EMBL:G03450
                     RHdb:RH53483
                     RHdb:RH3638
                     dbSTS:STS817
                     Identified using the e-PCR software (G. Schuler)"
     STS             123280. .123531
                     /note="matching EMBL:G05389
                     RHdb:RH53673
                     RHdb:RH3780
                     dbSTS:STS4997
                     Identified using the e-PCR software (G. Schuler)"
)RIGIN

  Query Match             28.0%;  Score 30;  DB 9;  Length 167398;
  Best Local Similarity   22.1%;  Pred. No. 1.3e+02;
  Matches   23;  Conservative   13;  Mismatches   68;  Indels    0;  Gaps    0;

)y          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              ||||      :  :|:|      |:   :::      | |   :| | | |      |::
)b     115884 AAAAAAATCATATTATGTATAATCTTTTTTTTTGGTGACAGGCTAGAGAAAAGGATGCTT 115825

)y         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUAGCA 106
                |          |||||:              ::| ||
)b     115824 GGGTGGCAATTGAAATAGAAATCCACTTTCAAATACTATTACCA 115781


!ESULT 4
)T001281
.OCUS       BT001281               1311 bp    mRNA    linear   INV 15-NOV-2002
)EFINITION  Drosophila melanogaster AT07872 full insert cDNA.
.CCESSION   BT001281
'ERSION     BT001281.1  GI:25009724
!EYWORDS    FLI_CDNA.
!OURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
!EFERENCE   1  (bases 1 to 1311)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Carlson,J.,
            Champe,M., Chavez,C., Dorsett,V., Dresnek,D., Farfan,D., Frise,E.,
            George,R., Gonzalez,M., Guarin,H., Kronmiller,B., Li,P., Liao,G.,
            Miranda,A., Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Patel,S., Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Rubin,G.M.
            and Celniker,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-NOV-2002) Berkeley Drosophila Genome Project,
            Lawrence Berkeley National Laboratory, One Cyclotron Road,
            Berkeley, CA 94720, USA
COMMENT     Sequence submitted by:
            Berkeley Drosophila Genome Project
             Lawrence Berkeley National Laboratory
            Berkeley, CA 94720
             This clone was sequenced as part of a high-throughput process to
            sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
            Science 2000). The sequence has been subjected to integrity checks
            for sequence accuracy, presence of a polyA tail and contiguity
            within 100 kb in the genome. Thus we believe the sequence to
            reflect accurately this particular cDNA clone. However, there are
            artifacts associated with the generation of cDNA clones that may
            have not been detected in our initial analyses such as internal
            priming, priming from contaminating genomic DNA, retained introns
            due to reverse transcription of unspliced precursor RNAs, and
            reverse transcriptase errors that result in single base changes.
            For further information about this sequence, including its location
            and relationship to other sequences, please visit our Web site
            (http://fruitfly.berkeley.edu) or send email to
            cdna@fruitfly.berkeley.edu.
FEATURES             Location/Qualifiers
     source          1. .1311
                     /organism="Drosophila melanogaster"
                     /mol_type="mRNA"
                     /db_xref="taxon:7227"
     gene            1. .1311
                     /gene="CG17138"
                     /db_xref="FLYBASE:FBgn0039021"
     CDS             36. .1229
                     /gene="CG17138"
                     /note="Longest ORF"
                     /codon_start=1
                     /product="AT07872p"
                     /protein_id="AAN71037.1"
                     /db_xref="GI:25009725"
                     /db_xref="FLYBASE:FBgn0039021"
                     /translation="MTQPRICFLVGALILSCTCVSCSKDKCAAIEHNPNSNTSYDFVP
                     HILHALAIHRPCVIGDPKCLCHVATIQRDLEPYVDKGITPEMMAQSKRLGTFYQIIRG
                     RIYRQQKCLHPKRCADVEDLLLDMASGVADLEFVLNVRDWPQVHFLSGLSGPVFSYSI
                     TNRHLDIMYPAWSFWTTTGPILQHYPHGVGRWDWMRKHLVARASELPWSAKRAIGFFR
                     GSRSSPERDSLVRLSQRRPDLVDAQYTILATDADPVEKMPLVEHCQFKYLFNFRGVAA
                     SFRLRHILLCRSLVLHVGDQWQEFFYSQLKPWVHYVPVASDADVDELAELILYLREHD
                     DLAEEIAERGQQFIWLHLRMEDVQCYWSKMLQEYAKLLTYKVQREPGLLEVSNKKAVQ
                     LYRGK"
ORIGIN

  Query Match             27.9%;  Score 29.8;  DB 3;  Length 1311;
  Best Local Similarity   26.5%;  Pred. No. 5.8e+02;
  Matches   22;  Conservative   12;  Mismatches   49;  Indels    0;  Gaps    0;

Qy          1 GGAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNC 60
              ||||||:     :  :|:       |:   :::      |||||  :| |||   :
Db       1221 GGAAAATAAACTTAGTATTTACAATCTTAGTTTTGTACCAAAAAATAAAAATCTGTATAA 1280

Qy         61 UUNAGNNNNNNNNNNNNAGAAA 83
              :: |             | |||
Db       1281 TTAAAAAAAAAAAAAAAAAAAA 1303


RESULT 5
AC117436/c
LOCUS       AC117436              133636 bp    DNA     linear   PRI 01-OCT-2002
DEFINITION  Homo sapiens 3 BAC RP11-69K16 (Roswell Park Cancer Institute Human
            BAC Library) complete sequence.
ACCESSION   AC117436
VERSION     AC117436.6  GI:23355769
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 133636)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D., Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstock,G. and Gibbs,R.
TITLE Direct Submission
JOURNAL Unpublished

REFERENCE 2 (bases 1 to 133636)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (10-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 133636)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (29-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 4 (bases 1 to 133636)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (30-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 5 (bases 1 to 133636)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (01-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT On Sep 30, 2002 this sequence version replaced gi:21553164.
INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

FEATURES	Location/Qualifiers
source	1. .133636
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="3"
	/clone="RP11-69K16"
STS	7. .197
	/standard_name="RH122785"
repeat_region	complement(513. .810)
	/rpt_family="AluSx"
repeat_region	833. .1131
	/rpt_family="AluJb"
repeat_region	1423. .1453
	/rpt_family="AT_rich"
repeat_region	1599. .1747
	/rpt_family="MER46C"
repeat_region	1794. .2232
	/rpt_family="MER9"
repeat_region	2233. .2253
	/rpt_family="(TG)n"
repeat_region	2254. .2310
	/rpt_family="MER9"
repeat_region	8037. .8312
	/rpt_family="HERVK_9"
repeat_region	8313. .8845
	/rpt_family="MER9"
repeat_region	complement(8886. .9207)
	/rpt_family="L1M4c"
repeat_region	9225. .9396
	/rpt_family="(TA)n"
repeat_region	complement(9405. .9825)
	/rpt_family="L1M4c"
repeat_region	complement(10094. .10457)
	/rpt_family="L1ME"
repeat_region	complement(10456. .10578)
	/rpt_family="L1MEc"
repeat_region	complement(10579. .10881)
	/rpt_family="AluJb"

```
     repeat_region   complement(10882. .11975)
                     /rpt_family="L1MEc"
     STS             11721. .11845
                     /standard_name="SHGC-77614"
     repeat_region   12033. .12114
                     /rpt_family="MER46C"
     repeat_region   12758. .13289
                     /rpt_family="L2"
     repeat_region   13532. .13569
                     /rpt_family="AT_rich"
     repeat_region   15616. .15945
                     /rpt_family="L1PB3"
     repeat_region   15946. .15984
                     /rpt_family="(TAAA)n"
     repeat_region   15985. .17070
                     /rpt_family="L1PB3"
     repeat_region   18445. .19144
                     /rpt_family="L1PB3"
     repeat_region   19145. .19166
                     /rpt_family="AT_rich"
     repeat_region   19292. .19417
                     /rpt_family="L1M4"
     repeat_region   19697. .20414
                     /rpt_family="MER21B"
     repeat_region   20465. .20575
                     /rpt_family="L2"
     repeat_region   20697. .21004
                     /rpt_family="AluY"
     repeat_region   22243. .22552
                     /rpt_family="AluSq"
     repeat_region   22553. .22717
                     /rpt_family="AluSg/x"
     repeat_region   23062. .23121
                     /rpt_family="(TA)n"
     repeat_region   23205. .23829
                     /rpt_family="L2"
     repeat_region   23799. .24083
                     /rpt_family="L2"
     repeat_region   24096. .24150
                     /rpt_family="L2"
     repeat_region   complement(24162. .24377)
                     /rpt_family="MIR"
     repeat_region   24378. .24447
                     /rpt_family="L2"
     repeat_region   24456. .24795

  Query Match             27.9%;  Score 29.8;  DB 9;  Length 133636;
  Best Local Similarity   23.4%;  Pred. No. 1.6e+02;
  Matches   25;  Conservative   12;  Mismatches   70;  Indels    0;  Gaps    0;

)y          1 GGAAAAUNNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNNAAAAANUANAAACAUNNNNNC 60
              |||||       : |: :|       :    ::      ||||  :  || ||:     |
)b      98804 GGAAACATCTATTAATTTGTGACTTTTTACATTGCCCAGAAAATTTGTAAGCATGGTCTC 98745

)y         61 UUNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAGCAG 107
              :  |               |||||:             :: | ||
)b      98744 TGAAATAGGCTAAAGAGAAGAAATGGAAAGGAAATGGCCTTTTGTAG 98698


(ESULT 6
.L359647
.OCUS       AL359647              86050 bp    DNA     linear   PRI 05-APR-2001
)EFINITION  Human DNA sequence from clone RP11-431P10 on chromosome 13,
            complete sequence.
.CCESSION   AL359647
'ERSION     AL359647.9  GI:13560015
(EYWORDS    HTG.
)OURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
(EFERENCE   1  (bases 1 to 86050)
  AUTHORS   Phillimore,B.
    TITLE     Direct Submission
    JOURNAL   Submitted (05-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
              CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
              requests: clonerequest@sanger.ac.uk
COMMENT       On Apr 6, 2001 this sequence version replaced gi:13273767.
              During sequence assembly data is compared from overlapping clones.
              Where differences are found these are annotated as variations
              together with a note of the overlapping clone name. Note that the
              variation annotation may not be found in the sequence submission
              corresponding to the overlapping clone, as we submit sequences with
              only a small overlap as described above.
              This sequence was finished as follows unless otherwise noted: all
              regions were either double-stranded or sequenced with an alternate
              chemistry or covered by high quality data (i.e., phred quality >=
              30); an attempt was made to resolve all sequencing problems, such
              as compressions and repeats; all regions were covered by at least
              one plasmid subclone or more than one M13 subclone; and the
              assembly was confirmed by restriction digest. The following
              abbreviations are used to associate primary accession numbers given
              in the feature table with their source databases: Em:, EMBL; Sw:,
              SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
              database can be found at
              http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
              was generated from part of bacterial clone contigs of human
              chromosome 13, constructed by the Sanger Centre Chromosome 13
              Mapping Group.  Further information can be found at
              http://www.sanger.ac.uk/HGP/Chr13
              RP11-431P10 is from the library RPCI-11.2 constructed by the group
              of Pieter de Jong. For further details see
              http://www.chori.org/bacpac/home.htm
              VECTOR: pBACe3.6
              IMPORTANT: This sequence is not the entire insert of clone
              RP11-431P10 It may be shorter because we sequence overlapping
              sections only once, except for a 100 base overlap.
              The true right end of clone RP11-431P10 is at 86050 in this
              sequence. The true left end of clone RP11-569O12 is at 48980 in
              this sequence. The true right end of clone RP11-600P1 is at 100 in
              this sequence.
FEATURES              Location/Qualifiers
     source           1. .86050
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /chromosome="13"
                      /clone="RP11-431P10"
                      /clone_lib="RPCI-11.2"
     misc_feature     110. .591
                      /note="match: GSS: Em:AQ621405"
     repeat_region    1507. .1694
                      /note="MER5A repeat: matches 1. .188 of consensus"
     repeat_region    1836. .2969
                      /note="L1M4 repeat: matches 2582. .3743 of consensus"
     repeat_region    2996. .3141
                      /note="L1MEc repeat: matches 1479. .1635 of consensus"
     repeat_region    3270. .4102
                      /note="L1MEc repeat: matches 1695. .2275 of consensus"
     repeat_region    4530. .6107
                      /note="L1PA15 repeat: matches 4541. .6157 of consensus"
     repeat_region    7573. .7622
                      /note="25 copies 2 mer gt 78% conserved"
     repeat_region    8343. .8653
                      /note="AluSx repeat: matches 1. .311 of consensus"
     repeat_region    9055. .9356
                      /note="MLT1E repeat: matches 251. .568 of consensus"
     misc_feature     complement(9811. .10118)
                      /note="match: GSS: Em:AQ092782"
     misc_feature     complement(10087. .10508)
                      /note="match: GSS: Em:AQ215557"
     misc_feature     10231. .10707
                      /note="match: GSS: Em:AQ728730"
     misc_feature     10260. .10779
                      /note="match: GSS: Em:AQ804059"
     repeat_region    11538. .11837
```

```
                     /note="AluY repeat: matches 3. .302 of consensus"
repeat_region        11838. .11873
                     /note="12 copies 3 mer taa 100% conserved"
misc_feature         14096. .14213
                     /note="match: STS: Em:G19980"
misc_feature         14097. .14420
                     /note="match: STS: Em:G06796"
repeat_region        14676. .14943
                     /note="MLT1A2 repeat: matches 4. .296 of consensus"
misc_feature         complement(14720. .15135)
                     /note="match: GSS: Em:AQ694000"
repeat_region        15052. .15118
                     /note="MLT1C repeat: matches 400. .466 of consensus"
repeat_region        15127. .15370
                     /note="L2 repeat: matches 790. .1041 of consensus"
repeat_region        15357. .15894
                     /note="L2 repeat: matches 2156. .2717 of consensus"
repeat_region        15929. .16389
                     /note="L2 repeat: matches 1057. .1516 of consensus"
repeat_region        16582. .16879
                     /note="AluJb repeat: matches 1. .312 of consensus"
repeat_region        17069. .17312
                     /note="MIR repeat: matches 5. .259 of consensus"
repeat_region        17590. .18043
                     /note="L1MA9 repeat: matches 5814. .6308 of consensus"
misc_feature         complement(17916. .18452)
                     /note="match: GSS: Em:AQ591830"
misc_feature         complement(18023. .18281)
                     /note="match: GSS: Em:AG019010"
misc_feature         complement(18036. .18247)
                     /note="match: GSS: Em:AQ375527"
repeat_region        18139. .18182
                     /note="22 copies 2 mer ta 79% conserved"
misc_feature         complement(18155. .18499)
                     /note="match: GSS: Em:AQ179736"
repeat_region        18296. .18710
                     /note="L1MA9 repeat: matches 5365. .5794 of consensus"
repeat_region        18711. .19930
                     /note="Tigger3b repeat: matches 3. .1231 of consensus"
repeat_region        19931. .19966
                     /note="L1MA9 repeat: matches 5329. .5365 of consensus"
repeat_region        19967. .20362
                     /note="MLT1A1 repeat: matches 3. .365 of consensus"
repeat_region        20363. .22158
                     /note="L1MA9 repeat: matches 3600. .5329 of consensus"
repeat_region        22167. .22346
                     /note="L2 repeat: matches 2021. .2197 of consensus"
misc_feature         complement(22629. .23020)
                     /note="match: GSS: Em:AQ133578"
repeat_region        23041. .23402
                     /note="THE1B repeat: matches 1. .364 of consensus"
repeat_region        23533. .23893
                     /note="L1PA15 repeat: matches 5791. .6157 of consensus"
repeat_region        24445. .24663
                     /note="AluJo repeat: matches 22. .244 of consensus"
repeat_region        26108. .26517
                     /note="MER57A repeat: matches 1. .426 of consensus"
repeat_region        26730. .26895
                     /note="MER53 repeat: matches 1. .188 of consensus"
repeat_region        27390. .28487
                     /note="L1PA3 repeat: matches 5044. .6146 of consensus"
repeat_region        28499. .29224
                     /note="L1PA3 repeat: matches 4286. .5014 of consensus"
repeat_region        29374. .29553
                     /note="MIR repeat: matches 20. .215 of consensus"
repeat_region        30291. .31294
                     /note="L1M2 repeat: matches 4712. .5713 of consensus"
repeat_region        31619. .31647
                     /note="MIR repeat: matches 196. .225 of consensus"
repeat_region        31648. .31944
                     /note="AluY repeat: matches 3. .299 of consensus"
repeat_region        31945. .32144
                     /note="MIR repeat: matches 6. .196 of consensus"
repeat_region        32300. .32365
                     /note="33 copies 2 mer cc 68% conserved"
repeat_region        32542. .33345
                     /note="L1PA5 repeat: matches 5342. .6145 of consensus"
repeat_region        33346. .33572
                     /note="L1PA5 repeat: matches 5107. .5333 of consensus"
repeat_region        33723. .33822
                     /note="L1MA9 repeat: matches 6199. .6300 of consensus"
misc_feature         complement(33849. .34245)
                     /note="match: GSS: Em:B32835"
repeat_region        34035. .34165
                     /note="L2 repeat: matches 2561. .2690 of consensus"
repeat_region        34745. .34809
                     /note="MIR repeat: matches 71. .134 of consensus"
repeat_region        34839. .37202
                     /note="TIGGER1 repeat: matches 1. .2414 of consensus"
misc_feature         40465. .40878
                     /note="match: GSS: Em:AQ109183"
repeat_region        41540. .41641
                     /note="L2 repeat: matches 2397. .2489 of consensus"
repeat_region        42086. .42218
                     /note="MER91A repeat: matches 46. .186 of consensus"
misc_feature         42489. .42816
                     /note="match: STS: Em:HSPH30H6"
repeat_region        44832. .45269
                     /note="MLT1C repeat: matches 6. .466 of consensus"
repeat_region        45290. .45388
                     /note="L2 repeat: matches 1662. .1748 of consensus"
repeat_region        45389. .45863
                     /note="MLT2FA repeat: matches 1. .448 of consensus"
repeat_region        45864. .46154
                     /note="L2 repeat: matches 1748. .2097 of consensus"
repeat_region        46192. .46414
                     /note="MLT1B repeat: matches 154. .390 of consensus"
repeat_region        46415. .46717
                     /note="AluSq repeat: matches 1. .307 of consensus"
repeat_region        46718. .46803
                     /note="MLT1B repeat: matches 112. .154 of consensus"
repeat_region        46804. .47085
                     /note="AluSg repeat: matches 1. .302 of consensus"
repeat_region        47086. .47191
                     /note="MLT1B repeat: matches 1. .112 of consensus"

  Query Match           27.7%;  Score 29.6;  DB 9;  Length 86050;
  Best Local Similarity  21.6%;  Pred. No. 2e+02;
  Matches   22;  Conservative   13;  Mismatches   67;  Indels    0;  Gaps    0;

Qy          1 GGAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNC 60
              | || |:      :  :|:|      |:   :::      | || :  |||||
Db      72475 GCAATATACCCCTACTATGAGCTTCCTTTCTTTTTCCCCTACAATTCCAAACAGTTCCTA 72534

Qy         61 UUNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUU 102
              :  ||           |||||:              ::
Db      72535 TGTAGAGTTCTTAGCAATAGAAATATGATTATCTATTATTTT 72576


RESULT 7
U67539
LOCUS       U67539                17492 bp    DNA     linear   BCT 28-JAN-1998
DEFINITION  Methanococcus jannaschii section 81 of 150 of the complete genome.
ACCESSION   U67539 L77117
VERSION     U67539.1  GI:1591619
KEYWORDS    .
SOURCE      Methanocaldococcus jannaschii
  ORGANISM  Methanocaldococcus jannaschii
            Archaea; Euryarchaeota; Methanococci; Methanococcales;
            Methanocaldococcaceae; Methanocaldococcus.
REFERENCE   1  (bases 1 to 17492)
  AUTHORS   Bult,C.J., White,O., Olsen,G.J., Zhou,L., Fleischmann,R.D.,
            Sutton,G.G., Blake,J.A., FitzGerald,L.M., Clayton,R.A.,
            Gocayne,J.D., Kerlavage,A.R., Dougherty,B.A., Tomb,J., Adams,M.D.,
            Reich,C.I., Overbeek,R., Kirkness,E.F., Weinstock,K.G.,
            Merrick,J.M., Glodek,A., Scott,J.D., Geoghagen,N.S., Weidman,J.F.,
```

Fuhrmann,J.L., Nguyen,D.T., Utterback,T., Kelley,J.M.,
Peterson,J.D., Sadow,P.W., Hanna,M.C., Cotton,M.D., Hurst,M.A.,
Roberts,K.M., Kaine,B.B., Borodovsky,M., Klenk,H.P., Fraser,C.M.,
Smith,H.O., Woese,C.R. and Venter,J.C.

TITLE Complete genome sequence of the methanogenic archaeon,
Methanococcus jannaschii

JOURNAL Science 273 (5278), 1058-1073 (1996)

MEDLINE 96337999

PUBMED 8688087

EFERENCE 2 (bases 1 to 17492)

AUTHORS Bult,C.J., White,O., Olsen,G.J., Zhou,L., Fleischmann,R.D.,
Sutton,G.G., Blake,J.A., FitzGerald,L.M., Clayton,R.A.,
Gocayne,J.D., Kerlavage,A.R., Dougherty,B.A., Tomb,J.-F.,
Adams,M.D., Reich,C.I., Overbeek,R., Kirkness,E.F., Weinstock,K.G.,
Merrick,J.M., Glodek,A., Scott,J.D., Geoghagen,N.S., Weidman,J.F.,
Fuhrmann,J.L., Nguyen,D.T., Utterback,T., Kelley,J.M.,
Peterson,J.D., Sadow,P.W., Hanna,M.C., Cotton,M.D., Hurst,M.A.,
Roberts,K.M., Kaine,B.B., Borodovsky,M., Klenk,H.P., Fraser,C.M.,
Smith,H.O., Woese,C.R. and Venter,J.C.

TITLE Direct Submission

JOURNAL Submitted (27-AUG-1996) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA

OMMENT On Oct 3, 1996 this sequence version replaced gi:1563994.

EATURES Location/Qualifiers

source 1. .17492
/organism="Methanocaldococcus jannaschii"
/mol_type="genomic DNA"
/db_xref="taxon:2190"

gene complement(87. .515)
/gene="MJ0952"

CDS complement(87. .515)
/gene="MJ0952"
/note="similar to SP:Q04493 PID:575709 percent identity:
26.72; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAB98957.1"
/db_xref="GI:1592316"
/translation="MENMAEDLRQKAMALEIYNQQLQMIQSEITSIRALKSEIMNSIK
TIENIKADEETLIPVGPGVFLKAKIVDDKALIGVKSDIYVEKSFNEVIEDLKKSVEDL
DKAEKEGMKKAEELAKAITALRKELQTEIQKAQQAQDKKQ"

gene 696. .998
/gene="MJ0953"

CDS 696. .998
/gene="MJ0953"
/note="similar to GB:AE000666 percent identity: 52.00;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAB98958.1"
/db_xref="GI:1499790"
/translation="MRLLGVIGYLAVLIKAICESWVDVVKRSINGEIHPQVIEIESII
NNPTGLVLLSWSITATPGTLVIDLIPEERKLKVAVISPRSREDIVPFEPYIKKIFD"

gene complement(1007. .2329)
/gene="MJ0954"

CDS complement(1007. .2329)
/gene="MJ0954"
/note="similar to GB:M59200 PID:150044 percent identity:
30.13; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="S-layer structural protein"
/protein_id="AAB98959.1"
/db_xref="GI:1591621"
/translation="MLRLQMMEGLIVKRTLLLILLLVISVSYALPIEPIIYVNKSTVD
YQNAKILMDNFYSSREININGDNVTIVINDIMYIPSIDELEIKNGDKNLIIKFDRDGN
KVKYKDIECIEYLNLKKGEEISLFNKSYIVEDITSNYVILKEKDGKEVLTNESFEYDG
YKVVVKLVSSDLNTIIVDIYKNEKVLDSPKLTKGKIYYMKGGTLGLMYENCTRIGKGY
RFTFRVYSTIKIEEGEDYPLDKEFKVKEISTDKIKLEYKNIDSLGNEIYLFNYTIIPE
KCYKDYVLFKVIKRKEKTVDVKDVAYIGDGIYAVKVNNTVHVFYKGKELKNHEKIYLG
SVDVYSSNPLNVNKDIILIGGPKVNKIVKELEDKGLLKVNISTNYPGNNRGIILKIKN
PYNDNNIYILAGSDRWGTKAAILVFLTKYNDEDTLMVEWDKGEIKIIK"

gene 2378. .3499
/gene="MJ0955"

CDS 2378. .3499
/gene="MJ0955"
/note="similar to SP:P06986 GB:U02071 GB:X03416 PID:41695
PID:41710 percent identity: 29.00; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="histidinol-phosphate aminotransferase (hisC)"
/protein_id="AAB98960.1"
/db_xref="GI:1591622"
/translation="MGDWMIENKVRDVVKKLKPYVPGKSKEEIARAYGIKPEDIIKLG
SNENPWGPSPKIKEKILDEIDKIHQYPEPVNPILMKELSKFLNVDEENIIVGGDGADE
IIDTIFRTFVDDGDEVIIPIPTFTQYRVSATIHNAKIKYAKYDKEKDFKLNVESVLNN
ITDKTKVIFLCTPNNPTGNIIENRDVERVINETDALVVIDHAYIEYAKKEYDWTQKAP
EYDNVIVLRTFSKVFGLAGMRVGYGVANKKIIDYMMRVKPIFSLTRLSQVCAITALRD
REFFERCVRDGIKSREMLYNGLKKFKDIKVYPSEANYLLVELKTMKAKEFCEELLKRG
VIVRDCTSFDGLGDNYVRVSIGTFEEVERFLKILEEIIS"

gene 3645. .4427
/gene="MJ0956"

CDS 3645. .4427
/gene="MJ0956"
/note="hypothetical protein; identified by GeneMark;
putative"
/codon_start=1
/transl_table=11
/product="M. jannaschii predicted coding region MJ0956"
/protein_id="AAB98969.1"
/db_xref="GI:1499793"
/translation="MHEDVWSLIKETNHKTILYSHGMKNRVLLGYLLKLRGKYNFDFE
AVTFISDRSPKWVREEIKWIMKANDVKHRFIECGKCKISLENNGHDYHGSCVVWEGIK
NFEGGVVASSIGEHIKDKERFLTWCNDYNLFPPFIHLNVSRERLVKEFESLDKNLISS
CYHACVFCPILYKLGKSLIKSEIFYKKKILTDNERKEHALNEFYDAMKEQDIGRMINS
VKIYYNKYIRELDEPYADYIKSRDVEMFDKLVNKCKEYLGVK"

gene 4427. .5479
/gene="MJ0957"

CDS 4427. .5479
/gene="MJ0957"
/note="hypothetical protein; identified by GeneMark;
putative"
/codon_start=1
/transl_table=11
/product="M. jannaschii predicted coding region MJ0957"
/protein_id="AAB98970.1"
/db_xref="GI:1499794"
/translation="MGDDIYIPPKPKPKPKPEYKYALWYFKYTNTYEKRFNGTVGD
LLNRGFNYAIALEKDEGSGTPESGNLREDGEKDGKEFGEFINSELSGIKYIAQIPYYK
RGMLEKLKNASKDKKQMEYYINHIYLVKRTLEYWKGWIDGVIESCDSNLVGFYWNFES
PGQVSWGFITDWEIAQLSTYIKQKSNELNRKLEFIWIPYINDIENPDNNDIKRLSKYF
DYVFVQPHYYIAWKYWCLWNYEKNVSEDIREYWKYQINRYNGYLTQGITKLIEVLNWI
KEIPNGYIEMEVDNKIDEYKYHDLINKACDYIKAREFLTGRDIWQIRAYYFDTNIENV
DKVRNGAYGIKGCKNW"

gene 5484. .6764
/gene="MJ0958"

CDS 5484. .6764
/gene="MJ0958"
/note="hypothetical protein; identified by GeneMark;
putative"
/codon_start=1
/transl_table=11
/product="M. jannaschii predicted coding region MJ0958"
/protein_id="AAB98971.1"
/db_xref="GI:1499795"
/translation="MKKFILFLIILLFSIYFLNVSSAEVCPFKDGFIIIYHDIGYDKL
FGYTYNDSEILYFNNKNLMDITPISDYYFPELYHNITYFECGSTNNLTILSFGYFYFP
KNGDGIVFLGVLVYTTKNNKINYTEKILWADVFYSVDDFDISPPACSPNEALLVYCYK
MKADYPENILVLINNTNITELKKFEDDDYYTTFQHYIYQPIFIFTTYDSKAKKFYILD
GRLSNTSFPLYSYYGGKIHFEDNITLPKYENISWECMDFYSINGTLYIVMKKLNCSNI
NYAGWYEDYLNQKPYLLIWKNKTIKIISNCSNPYYFVKIKGGEIPIEKSYLKKIFGNK
YESIRITDLAYNNGILLIETNENHKLHYYIVKNNSIEEIKLKNIIKLYKKKHSLWDDI
KKELEPKIYWVIHNWYIIVLIIAGLLWMAILWKK"

```
    gene            complement(6788. .7945)
                    /gene="MJ0959"
    CDS             complement(6788. .7945)
                    /gene="MJ0959"
                    /note="similar to PID:704449 GB:AE000666 percent identity:
                    54.69; identified by sequence similarity; putative"
                    /codon_start=1
                    /transl_table=11
                    /product="aspartate aminotransferase (aspC)"
                    /protein_id="AAB98961.1"
                    /db_xref="GI:1591623"
                    /translation="MKIDAVKKLLMIPGPTMVPPEVLNAMALPVIGHRTKDYSNLLED
                    TIEKLKKVFITENDTFLITGSGTAAMDMAISNIIKRGDKVLNIVTGNFGERFANIVKA
                    YKGEAIRLDVEWGDMAEPEAVKEILDKYDDIKAVTVVHNETSTGARNPIKEIGEVVKD
                    YDALYIVDTVSSLGGDYVNVDKFHIDICVTGSQKCLAAPPGLAAITVSEKAWEVIKKN
                    DDKVGFYLDLLAYKKYYEEKKQTPYTPSVNLTYALNVALDLVLEEGIENRVKRHERLA
                    KATRAGLEAMGIELFAKERARSVTVTSAKYPEGIEDSKFRGILSNKYNIVVAGGQKHL
                    AGKIFRIGHMGICGEKEVLATLACVELALKELGFEVKESGVEVAKEVLLKE"
    gene            complement(7990. .8643)
                    /gene="MJ0960"
    CDS             complement(7990. .8643)
                    /gene="MJ0960"
                    /note="similar to SP:P19669 PID:853766 GB:AL009126 percent
                    identity: 59.02; identified by sequence similarity;
                    putative"
                    /codon_start=1
                    /transl_table=11
                    /product="transaldolase"
                    /protein_id="AAB98962.1"
                    /db_xref="GI:1591624"
                    /translation="MKFFLDTANVEEIKKYAELGLVDGVTTNPTLVAKEGRDFYEVVK
                    EICEIVEGPVSAEVISTDAEGMVKEARELAKLADNIVIKIPMTKDGMKAVKILSAEGI
                    KTNVTLVFSPLQALVAAKAGATYVSPFVGRLDDIGHVGMKLIEDVVKIYKNYDIKTEV
                    IVASVRHPWHVLEAAKIGADIATMPPAVMDKLFNHPLTDIGLERFLKDWDEYLKSRK"
    gene            8860. .11148
                    /gene="MJ0961"
    CDS             8860. .11148
                    /gene="MJ0961"
                    /note="similar to SP:P30665 PID:1314093 PID:3482
```

```
  Query Match             27.5%;  Score 29.4;  DB 1;  Length 17492;
  Best Local Similarity   24.8%;  Pred. No. 3.5e+02;
  Matches   26;  Conservative   10;  Mismatches   69;  Indels    0;  Gaps    0;

Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              ||||:     : | | |       |:   :::      ||||| :| |||||:      :
Db      15270 AAAATTAGTATTAGAAGCTATAACTTATTTTAATGCCAAAAAATATAAACATTTCTGGTG 15329

Qy         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAGCAG 107
                |                ||  |:                | ||
Db      15330 AAATAATGACAGGCAAAGTTATCTTAGTAGGAGCAGGACCGGGAG 15374


RESULT 8
AR271569_05/c
WPCOMMENT
  Sequence split into 17 fragments  LOCUS AR271569 Accession AR271569
     Fragment Name      Begin     End
     AR271569_00            1    110000
     AR271569_01       100001    210000
     AR271569_02       200001    310000
     AR271569_03       300001    410000
     AR271569_04       400001    510000
     AR271569_05       500001    610000
     AR271569_06       600001    710000
     AR271569_07       700001    810000
     AR271569_08       800001    910000
     AR271569_09       900001   1010000
     AR271569_10      1000001   1110000
     AR271569_11      1100001   1210000
     AR271569_12      1200001   1310000
     AR271569_13      1300001   1410000
     AR271569_14      1400001   1510000
     AR271569_15      1500001   1610000
     AR271569_16      1600001   1664976
Continuation (6 of 17) of AR271569 from base 500001 (AR271569 Sequence 1 from patent US

  Query Match             27.5%;  Score 29.4;  DB 6;  Length 110000;
  Best Local Similarity   24.8%;  Pred. No. 2.1e+02;
  Matches   26;  Conservative   10;  Mismatches   69;  Indels    0;  Gaps    0;

Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              ||||:     : | | |       |:   :::      ||||| :| |||||:      :
Db      73299 AAAATTAGTATTAGAAGCTATAACTTATTTTAATGCCAAAAAATATAAACATTTCTGGTG 73240

Qy         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAGCAG 107
                |                ||  |:                | ||
Db      73239 AAATAATGACAGGCAAAGTTATCTTAGTAGGAGCAGGACCGGGAG 73195


RESULT 9
AC015173/c
LOCUS       AC015173               60471 bp    DNA     linear   HTG 16-NOV-1999
DEFINITION  Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***.
ACCESSION   AC015173
VERSION     AC015173.1  GI:6436162
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 60471)
  AUTHORS   Adams,M. and Venter,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
COMMENT     This sequence was identified as CDM:10210450 by the submitter.
            For further information on this sequence e-mail to fly@celera.com.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source          1. .60471
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
ORIGIN

  Query Match             27.3%;  Score 29.2;  DB 2;  Length 60471;
  Best Local Similarity   25.0%;  Pred. No. 2.7e+02;
  Matches   21;  Conservative   13;  Mismatches   50;  Indels    0;  Gaps    0;

Qy          1 GGAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNC 60
              ||||||:     :  :|:       |:   :::      ||||| :| |||  :
Db      15041 GGAAAATAAACTTAGTATTTACAATCTTAGTTTTGTACCAAAAAATAAAAATCTGTATAA 14982

Qy         61 UUNAGNNNNNNNNNNNNNAGAAAU 84
              :: |               |||:
Db      14981 TTAATTTTAAATTCAAATTTAAAT 14958


RESULT 10
AC009846/c
LOCUS       AC009846              164038 bp    DNA     linear   INV 17-FEB-2001
DEFINITION  Drosophila melanogaster, chromosome 3R, region 94C-94D, BAC clone
            BACR23F10, complete sequence.
ACCESSION   AC009846
VERSION     AC009846.9  GI:12957616
KEYWORDS    HTG.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
```

```
EFERENCE   1  (bases 1 to 164038)
 AUTHORS   Celniker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
           Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,
           Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A.,
           Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,
           Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
           Ferriera,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A.,
           Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
           Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
           McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
           Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B.,
           Phouanenavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
           Stapleton,M., Strong,R., Svirskas,R., Tector,C., Williams,S.M.,
           Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
  TITLE    Sequencing of Drosophila chromosome 3R, region 94C-94D
 JOURNAL   Unpublished
EFERENCE   2  (bases 1 to 164038)
 AUTHORS   Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
           Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
           Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
           Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
           Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
           Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
           Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,
           Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
           Rubin,G.M.
  TITLE    Direct Submission
 JOURNAL   Submitted (03-SEP-1999) Drosophila Genome Center, Lawrence Berkeley
           Laboratory, MS 64-121, Berkeley, CA 94720, USA
OMMENT     On Feb 17, 2001 this sequence version replaced gi:6957584.
           Sequence submitted by:
           Berkeley Drosophila Genome Project
           Lawrence Berkeley National Laboratory, MS 64-121
           Berkeley, CA 94720
           This sequence was assembled using end sequences from a whole genome
           shotgun and from subclones of this BAC and its neighboring clones.
           For further information about this sequence, including its location
           and relationship to other sequences, please visit our sequence
           archive Web site (http://www.fruitfly.org/sequence/) or send email
           to bdgp@fruitfly.berkeley.edu.
EATURES              Location/Qualifiers
     source          1. .164038
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /strain="y; cn bw sp"
                     /db_xref="taxon:7227"
                     /chromosome="3R"
                     /map="94C-94D"
                     /clone="BACR23F10 (D1097)"
                     /clone_lib="RPCI-98 (Roswell Park Cancer Institute
                     Drosophila melanogaster BAC library, partial EcoRI in
                     pBACe3.6)"
)RIGIN

  Query Match             27.3%;  Score 29.2;  DB 3;  Length 164038;
  Best Local Similarity   25.0%;  Pred. No. 2e+02;
  Matches   21;  Conservative   13;  Mismatches   50;  Indels    0;  Gaps    0;

)y          1 GGAAAAAUNNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNNAAAAANUANAAACAUNNNNNC 60
              ||||||:     : :|:       |:   :::      |||||:|  |||  :
)b     115236 GGAAAATAAACTTAGTATTTACAATCTTAGTTTTGTACCAAAAAATAAAAATCTGTATAA 115177

)y         61 UUNAGNNNNNNNNNNNNNAGAAAU 84
              :: |              |||:
)b     115176 TTAATTTTAAATTCAAATTTAAAT 115153


(ESULT 11
\E003740/c
.OCUS       AE003740               234378 bp    DNA     linear   INV 14-FEB-2003
)EFINITION  Drosophila melanogaster chromosome 3R, section 78 of 118 of the
            complete sequence.
\CCESSION   AE003740 AE002708 AE014297
VERSION     AE003740.3  GI:23171988
KEYWORDS    .
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 234378)
  AUTHORS   Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
            Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F.,
            George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,
            Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X.,
            Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D.,
            Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor,G.L.,
            Abril,J.F., Agbayani,A., An,H.J., Andrews-Pfannkoch,C., Baldwin,D.,
            Ballew,R.M., Basu,A., Baxendale,J., Bayraktaroglu,L., Beasley,E.M.,
            Beeson,K.Y., Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S.,
            Borkova,D., Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P.,
            Burtis,K.C., Busam,D.A., Butler,H., Cadieu,E., Center,A.,
            Chandra,I., Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B.,
            Davies,P., de Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I.,
            Dietz,S.M., Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S.,
            Dunkov,B.C., Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C.,
            Ferriera,S., Fleischmann,W., Fosler,C., Gabrielian,A.E., Garg,N.S.,
            Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,
            Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J.,
            Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J.,
            Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z.,
            Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,
            Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A.,
            Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C.,
            McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C.,
            Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L.,
            Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K.,
            Nusskern,D.R., Pacleb,J.M., Palazzolo,M., Pittman,G.S., Pan,S.,
            Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K.,
            Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I.,
            Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C.,
            Stapleton,M., Strong,R., Sun,E., Svirskas,R., Tector,C., Turner,R.,
            Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A.,
            Weinstock,G.M., Weissenbach,J., Williams,S.M., WoodageT,
            Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F.,
            Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H.,
            Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,
            Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.
  TITLE     The genome sequence of Drosophila melanogaster
 JOURNAL    Science 287 (5461), 2185-2195 (2000)
 MEDLINE    20196006
  PUBMED    10731132
REFERENCE   2  (bases 1 to 234378)
  AUTHORS   Celniker,S.E., Adams,M.D., Kronmiller,B., Wan,K.H., Holt,R.A.,
            Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C., Rogers,Y.,
            Banzon,J., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A.,
            Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,
            Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
            Ferriera,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A.,
            Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
            Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
            McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
            Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B.,
            Phouanenavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
            Stapleton,M., Strong,R., Svirskas,R., Tector,C., Tyler,D.,
            Williams,S.M., Zaveri,J.S., Smith,H.O., Venter,J.C. and Rubin,G.M.
  TITLE     Sequencing of Drosophila melanogaster genome
 JOURNAL    Unpublished
REFERENCE   3  (bases 1 to 234378)
  AUTHORS   Misra,S., Crosby,M.A., Matthews,B.B., Bayraktaroglu,L.,
            Campbell,K., Hradecky,P., Huang,Y., Kaminker,J.S., Prochnik,S.E.,
            Smith,C.D., Tupy,J.L., Bergman,C.M., Berman,B.P., Carlson,J.W.,
            Celniker,S.E., Clamp,M.E., Drysdale,R.A., Emmert,D., Frise,E., de
            Grey,A.D.N.J., Harris,N.L., Kronmiller,B., Marshall,B.,
            Millburn,G.H., Richter,J., Russo,S., Searle,S.M.J., Smith,E.,
            Shu,S., Smutniak,F., Whitfield,E.J., Ashburner,M., Gelbart,W.M.,
```

```
            Rubin,G.M., Mungall,C.J. and Lewis,S.E.
  TITLE     Annotation of Drosophila melanogaster genome
  JOURNAL   Unpublished
REFERENCE   4  (bases 1 to 234378)
  AUTHORS   Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
REFERENCE   5  (bases 1 to 234378)
  AUTHORS   .
  CONSRTM   FlyBase
  TITLE     Direct Submission
  JOURNAL   Submitted (06-SEP-2002) University of California Berkeley, 539 Life
            Sciences Addition, Berkeley, CA 94720, USA
REFERENCE   6  (bases 1 to 234378)
  AUTHORS   .
  CONSRTM   FlyBase
  TITLE     Direct Submission
  JOURNAL   Submitted (13-FEB-2003) University of California Berkeley, 539 Life
            Sciences Addition, Berkeley, CA 94720, USA
  REMARK    Sequence update by submitter
COMMENT     On Sep 18, 2002 this sequence version replaced gi:10726705.
FEATURES             Location/Qualifiers
     source          1. .234378
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /chromosome="3R"
                     /note="genotype: y[1]; cn[1] bw[1] sp[1]; Rh6[1]"
     gene            10750. .13380
                     /locus_tag="CG17622"
                     /note="last curated on Tue Jan 22 00:31:41 PST 2002"
                     /map="94B10-94B10"
                     /db_xref="FLYBASE:FBgn0038998"
     mRNA            join(10750. .10772,11010. .11215,12157. .12598,12859. .13185,
                     13311. .13380)
                     /locus_tag="CG17622"
                     /product="CG17622-RA"
                     /db_xref="FLYBASE:FBgn0038998"
     CDS             join(10750. .10772,11010. .11215,12157. .12598,12859. .13185,
                     13311. .13380)
                     /locus_tag="CG17622"
                     /note="CG17622 gene product"
                     /codon_start=1
                     /product="CG17622-PA"
                     /protein_id="AAF56036.1"
                     /db_xref="GI:7300896"
                     /db_xref="FLYBASE:FBgn0038998"
                     /translation="MVPFFRSVAQQSAAAAASGAGGAGGAGGGGGGSGSQSRPGSRKL
                     PPPLLRSRTLPAIIVPGLPVVSLHTDKQTFQLDERLLGSKSRSGSASGHRWSLLTRAT
                     NSSSMCSNNNPASSSSSSSNNNNNLHMANNNNTLMVKSLNLPKDDSNLVYRRKSSGST
                     PHQHQNAHHHHQQHHHLQQQQQQQQQHQQQQLQQQQMLQHLGDSSLFQLSDDFECHAD
                     GSLLLRIPAPHVVAARAYRLAARKRAAASSEASTPMTREHTLHEMPLAESACSGSATG
                     SGTTFTRLAKLLHQSSTLAGRSNLQPPPMADNAVQSLGRGLGLGLGEEAPRRLSWESW
                     KADLGRVSLAAPPNGRPHEVG"
     gene            18885. .19523
                     /locus_tag="CG17623"
                     /note="last curated on Tue Jan 22 00:33:38 PST 2002"
                     /map="94B11-94B11"
                     /db_xref="FLYBASE:FBgn0038999"
     mRNA            join(18885. .18931,19007. .19523)
                     /locus_tag="CG17623"
                     /product="CG17623-RA"
                     /db_xref="FLYBASE:FBgn0038999"
     CDS             join(18885. .18931,19007. .19523)
                     /locus_tag="CG17623"
                     /note="CG17623 gene product"
                     /codon_start=1
                     /product="CG17623-PA"
                     /protein_id="AAF56037.2"
                     /db_xref="GI:23171989"
                     /db_xref="FLYBASE:FBgn0038999"
                     /translation="MPSALHLRESPPESIERKDSSALQRSASIDSFAEIVFSESPRPS
                     LAVTGPGACASASGGPSPFSKRPSASSLYSTSTSSSFASQMQAQLNVNYGDSALGSGS
                     ATGGPGSGHHHGASGGGNGGSSASSSRRESMLSPSSTRRSKLTRIINGECGVNRMRNP
                     KIYCQAMRALCMDSTIQTSYLYTLLTI"
     gene            28305. .36625
                     /locus_tag="CG6954"
                     /note="last curated on Tue Jan 22 00:49:10 PST 2002"
                     /map="94C1-94C1"
                     /db_xref="FLYBASE:FBgn0039000"
     mRNA            join(28305. .28975,30084. .30334,31045. .31599,31729. .31807,
                     31872. .31975,32036. .32173,32240. .32464,32702. .32908,
                     32966. .33176,33240. .36625)
                     /locus_tag="CG6954"
                     /product="CG6954-RA"
                     /db_xref="FLYBASE:FBgn0039000"
     CDS             join(28305. .28975,30084. .30334,31045. .31599,31729. .31807,
                     31872. .31975,32036. .32173,32240. .32464,32702. .32908,
                     32966. .33176,33240. .35715)
                     /locus_tag="CG6954"
                     /note="CG6954 gene product"
                     /codon_start=1
                     /product="CG6954-PA"
                     /protein_id="AAF56038.1"
                     /db_xref="GI:7300898"
                     /db_xref="FLYBASE:FBgn0039000"
                     /translation="MERRKALSSYAKFKDQQEAAAAVTPQKDNESDDSLPALRRSHSV
                     RSSLRRLKNKLHSPALSMQSPFKLRTQLHMRNRSSRGSATLDKSKAAKALRMLGDLDA
                     VAVLTTKVKGKNEFVPLKRAMNSDSCDHDELGIERLMSSLPVGAKLPRKACKLLQIPE
                     SYCRQAEVAPKATLDSDLEKTFSGEESSAVLADITRHAQQGIYGTTRMRTATIRKPMP
                     YLNSRRCYHSDSTKLPNMPQKLPMSSCLRSPAVPVRQKSFHTLNELTRFCQLKVSPHR
                     STATATSTSSSPSSSPAAAQQQRRHSHYAIDDNILPQDKRDGDIDSVRLREKVVNPSG
                     GNNNNQDNYPTYYPAPLLSTSRWSEQLTQHAAVAKIRTAISCHSQDLREMEFLLPHSR
                     TDSQSNSQSQLCISRHSQPKNAKLRDDHDETGSSELDSPPLGDENASANGQLSVHRQP
                     VQLRSKIPGNTPHPRASKMSKKQLKLAQAQLDKLTQNNLHLHALFSAVEHGHLEKART
                     ILESTDVDVNSINNDGLSALDLAVLSNNRSMTRMLLQHGAVEGSQFSVDTIGTKLNGL
                     LKDAESRIHDLSGPEGLCPPVFASRPSISSIIIGNSSASVTGCTGSEVEKQIGIWERR
                     VKGLRRLQLGWDQARPPDAPASVVVDVTGDNSISVQILEPFEGAIGTKFKVQWSTRAD
                     FNNVVGESELLEWISFHGTMGAQCHISGLTQGRRYFLRAACGNVKGWGTYRTSVPASV
                     VPSTWRDLDNREDRFVGRHRILDNLFTAVRLARPADVSELTLDPASAQRRNPKKKTTI
                     KQLFSVTTKFQKTLRRGIYFSCIIHCDDKVLVTSEDFPPVIEVDESYPSALHMDYYWL
                     MKVACTWEDVKSLRSDMERNLTSAVHFRTKLLSAVCQMQSALGITDLGQLYYKPLRDA
                     QGTVVLTCVQSVKSQKAVSILNSRWVPVSKLQKKLGALHEDYTINELLISSIGDQLHY
                     QQAALQRLEPGLYLGYLKMQCSIDQIQVVVPVKTPNVLPHCKVRENSHITAEEWQVLH
                     RYSSDPLRLPLDFSAQGGDGASGATATTEVQRLFLYDLTNAMHKLFASMNIKIADATT
                     HRLYDVEVIEHSPDISFLVVCPSAESSCAVPGQSELLLQRDDLASLSIQAFEMIHLRT
                     YQPAIIQKYARLSCILELDTALATHSLREAFSSSELQAAKERLATLQELSASLTIVWK
                     SVRWLMDVVAYARNKNAQPSLAMREILDFAQQRQDEAVATSAGGSANKQLLQLPIRES
                     KFSKTGQGRGSWPGPGTSEDQSQNKPEHSKSEQNLELSAITPVEPASNQVPTQQQQPQ
                     QQSQQQQLLQVSNISEYAGSICSEVSFRKNSGDSMSSTYTSRSFYSAVDSASDGNSTN
                     SVFAIPPSRSDDTLADALRHSQAVAAQRKRTSSNIASHTNPLITVHSSSSAPYLAGSS
                     VSLRSGNGAFATDLEHKPLKPATKAASSANLREGGPYLKSTLAELRAVGGEEPVASTS
```

```
  Query Match             27.3%;  Score 29.2;  DB 3;  Length 234378;
  Best Local Similarity   25.0%;  Pred. No. 1.9e+02;
  Matches   21;  Conservative   13;  Mismatches   50;  Indels    0;  Gaps    0;

Qy          1 GGAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAAANUANAAACAUNNNNNC 60
              ||||||:     :  :|:       |:   :::      |||||  :| |||  :
Db     101887 GGAAAATAAACTTAGTATTTACAATCTTAGTTTTGTACCAAAAAATAAAAATCTGTATAA 101828

Qy         61 UUNAGNNNNNNNNNNNNNAGAAAU 84
              :: |               |||:
Db     101827 TTAATTTTAAATTCAAATTTAAAT 101804


RESULT 12
AC079951
LOCUS       AC079951               191942 bp    DNA     linear   PRI 15-MAR-2002
DEFINITION  Homo sapiens 12q BAC RP11-346L11 (Roswell Park Cancer Institute
            Human BAC Library) complete sequence.
ACCESSION   AC079951
VERSION     AC079951.35  GI:15055247
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
EFERENCE 1 (bases 1 to 191942)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M., Neal,D., Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shim,C., Shooshtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kucherlapati,R., Weinstock,G. and Gibbs,R.
TITLE Direct Submission
JOURNAL Unpublished
EFERENCE 2 (bases 1 to 191942)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (20-SEP-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
EFERENCE 3 (bases 1 to 191942)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (01-AUG-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
EFERENCE 4 (bases 1 to 191942)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (09-AUG-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
EFERENCE 5 (bases 1 to 191942)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (15-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
OMMENT On Aug 1, 2001 this sequence version replaced gi:14861675.
INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

QUALSTAT-REPORT.

FEATURES	Location/Qualifiers
source	1. .191942
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="12"
	/clone="RP11-346L11"
repeat_region	582. .883
	/rpt_family="AluSg"
repeat_region	1600. .1799
	/rpt_family="MER46A"
repeat_region	1906. .2212
	/rpt_family="AluSq"
repeat_region	3606. .3915
	/rpt_family="AluSc"
repeat_region	4874. .4917
	/rpt_family="AT_rich"
repeat_region	complement(5619. .5883)
	/rpt_family="AluSq"
repeat_region	6028. .6335
	/rpt_family="AluSx"
repeat_region	6341. .6774
	/rpt_family="L1MCb"
repeat_region	6918. .7034
	/rpt_family="L1M4"
repeat_region	7038. .7264
	/rpt_family="L1MA2"
repeat_region	complement(7337. .7633)
	/rpt_family="HUERS-P3b"
repeat_region	complement(7659. .11980)
	/rpt_family="L1PA5"
STS	9027. .9204
	/standard_name="61535"
repeat_region	complement(11981. .13272)
	/rpt_family="SVA"
repeat_region	complement(13273. .15103)

```
                     /rpt_family="L1PA5"
     repeat_region   complement(16729. .17246)
                     /rpt_family="HUERS-P3b"
     repeat_region   complement(17247. .17899)
                     /rpt_family="LTR9B"
     repeat_region   19794. .20065
                     /rpt_family="AluJb"
     repeat_region   21526. .21568
                     /rpt_family="MIR"
     repeat_region   21589. .21885
                     /rpt_family="AluSx"
     repeat_region   22537. .22570
                     /rpt_family="(TAAA)n"
     repeat_region   22606. .22691
                     /rpt_family="MIR"
     repeat_region   24218. .24508
                     /rpt_family="AluSc"
     repeat_region   24566. .24738
                     /rpt_family="FAM"
     repeat_region   24925. .24990
                     /rpt_family="L2"
     repeat_region   25052. .25349
                     /rpt_family="AluSq"
     repeat_region   25352. .25649
                     /rpt_family="AluSx"
     repeat_region   26062. .26197
                     /rpt_family="L1MA9"
     repeat_region   26368. .26637
                     /rpt_family="AluSq"
     repeat_region   complement(27254. .27352)
                     /rpt_family="MIR"
     repeat_region   complement(27925. .28222)
                     /rpt_family="AluSq"
     misc_feature    complement(28240. .31468)
                     /note="Region: HSLACTG Human alpha-lactalbumin gene"
     STS             28464. .28721
                     /standard_name="68443"
     STS             28558. .28755
                     /standard_name="50357"
     repeat_region   30200. .30485
                     /rpt_family="AluSq"
     repeat_region   30486. .30515
                     /rpt_family="(CATA)n"
     repeat_region   31469. .31773
                     /rpt_family="AluJb"
     repeat_region   complement(32145. .32452)
                     /rpt_family="AluSx"

  Query Match             27.1%;  Score 29;  DB 9;  Length 191942;
  Best Local Similarity   23.3%;  Pred. No. 2.2e+02;
  Matches   24;  Conservative   11;  Mismatches   68;  Indels    0;  Gaps    0;

Qy          1 GGAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNC 60
              ||| ||:     : |  :      |:   :::     ||| | :  |||||:
Db     155959 GGAGAATGAAATTAACTTTTTCTGTCTGGATTTAGGAGAAAACAATGGAAACATTTTAAA 156018

Qy         61 UUNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUA 103
                ||             |||||               ::|
Db     156019 ACAAGTTTTATATCTGACAGAAAAAAAAGGAATGATTCCATTA 156061


RESULT 13
CEC01B9/c
LOCUS       CEC01B9                25233 bp    DNA     linear   INV 03-DEC-2003
DEFINITION  Caenorhabditis elegans cosmid C01B9, complete sequence.
ACCESSION   Z93373
VERSION     Z93373.1  GI:3217060
KEYWORDS    HTG.
SOURCE      Caenorhabditis elegans
  ORGANISM  Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1
  AUTHORS   none.
  TITLE     Genome sequence of the nematode C. elegans: a platform for
            investigating biology. The C. elegans Sequencing Consortium
  JOURNAL   Science 282 (5396), 2012-2018 (1998)
  MEDLINE   99069613
   PUBMED   9851916
  REMARK    The C.elegans Sequencing Consortium.
REFERENCE   2  (bases 1 to 25233)
  AUTHORS   Percy,C.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAR-1997) Nematode Sequencing Project, Sanger
            Institute, Hinxton, Cambridge CB10 1SA, England and Department of
            Genetics, Washington University, St. Louis, MO 63110, USA. E-mail:
            jes@sanger.ac.uk or rw@nematode.wustl.edu
COMMENT     On Jun 13, 1998 this sequence version replaced gi:1913957.
            Coding sequences below are predicted from computer analysis, using
            predictions from Genefinder (P. Green, U. Washington), and other
            available information.
            Current sequence finishing criteria for the C. elegans genome
            sequencing consortium are that all bases are either sequenced
            unambiguously on both strands, or on a single strand with both a
            dye primer and dye terminator reaction, from distinct subclones.
            Exceptions are indicated by an explicit note.
            For a graphical representation of this sequence and its analysis
            see:- http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?
            name=C01B9
            IMPORTANT:  This sequence is NOT necessarily the entire insert of
            the specified clone.  It may be shorter because we only sequence
            overlapping sections once, or longer because we arrange for a small
            overlap between neighbouring submissions.
            This sequence is the entire insert of clone C01B9. The true left
            end of clone ZK131 is at 18285 in this sequence. The true right end
            of clone R06B9 is at 100 in this sequence. The start of this
            sequence (1. .100) overlaps with the end of sequence CER06B9.
            The end of this sequence (24951. .25233) overlaps with the start of
            sequence CEZK131.
FEATURES             Location/Qualifiers
     source          1. .25233
                     /organism="Caenorhabditis elegans"
                     /mol_type="genomic DNA"
                     /strain="Bristol N2"
                     /db_xref="taxon:6239"
                     /clone="C01B9"
ORIGIN

  Query Match             26.7%;  Score 28.6;  DB 3;  Length 25233;
  Best Local Similarity   21.8%;  Pred. No. 4.9e+02;
  Matches   22;  Conservative   12;  Mismatches   67;  Indels    0;  Gaps    0;

Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              | ||     : |: :      |:   :::     |||  :| ||| |:     |:
Db      23035 ATAAGCTTGCTAATTTTTTCTGGCTTTTTTTTAAATAGAAATATAGAAAAATATCGGCTG 22976

Qy         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUA 103
                ||           |||||               ::|
Db      22975 GAGTTTAAATTTTCTGAGAAAAATTTTTGGCGGGAATTTTA 22935


RESULT 14
PFMAL3P4/c
LOCUS       PFMAL3P4               113880 bp    DNA     linear   INV 29-JAN-2003
DEFINITION  Plasmodium falciparum MAL3P4.
ACCESSION   AL008970 AL139179 AL844502
VERSION     AL008970.3  GI:7672212
KEYWORDS    .
SOURCE      Plasmodium falciparum 3D7
  ORGANISM  Plasmodium falciparum 3D7
            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE   1  (bases 1 to 113880)
  AUTHORS   Bowman,S., Lawson,D., Basham,D., Brown,D., Chillingworth,T.,
            Churcher,C.M., Craig,A., Davies,R.M., Devlin,K., Feltwell,T.,
            Gentles,S., Gwilliam,R., Hamlin,N., Harris,D., Holroyd,S.,
            Hornsby,T., Horrocks,P., Jagels,K., Jassal,B., Kyes,S., McLean,J.,
```

Moule,S., Mungall,K., Murphy,L., Oliver,K., Quail,M.A., Rajandream,M.-A., Rutter,S., Skelton,J., Squares,R., Squares,S., Sulston,J.E., Whitehead,S., Woodward,J.R., Newbold,C. and Barrell,B.G.
TITLE The complete nucleotide sequence of chromosome 3 of Plasmodium falciparum
JOURNAL Nature 400 (6744), 532-538 (1999)
MEDLINE 99376085
PUBMED 10448855
REFERENCE 2
AUTHORS Hall,N., Pain,A., Berriman,M., Churcher,C., Harris,B., Harris,D., Mungall,K., Bowman,S., Atkin,R., Baker,S., Barron,A., Brooks,K., Buckee,C.O., Burrows,C., Cherevach,I., Chillingworth,C., Chillingworth,T., Christodoulou,Z., Clark,L., Clark,R., Corton,C., Cronin,A., Davies,R., Davis,P., Dear,P., Dearden,F., Doggett,J., Feltwell,T., Goble,A., Goodhead,I., Gwilliam,R., Hamlin,N., Hance,Z., Harper,D., Hauser,H., Hornsby,T., Holroyd,S., Horrocks,P., Humphray,S., Jagels,K., James,K.D., Johnson,D., Kerhornou,A., Knights,A., Konfortov,B., Kyes,S., Larke,N., Lawson,D., Lennard,N., Line,A., Maddison,M., Mclean,J., Mooney,P., Moule,S., Murphy,L., Oliver,K., Ormond,D., Price,C., Quail,M.A., Rabbinowitsch,E., Rajandream,M.A., Rutter,S., Rutherford,K.M., Sanders,M., Simmonds,M., Seeger,K., Sharp,S., Smith,R., Squares,R., Squares,S., Stevens,K., Taylor,K., Tivey,A., Unwin,L., Whitehead,S., Woodward,J., Sulston,J.E., Craig,A., Newbold,C. and Barrell,B.G.
TITLE Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13
JOURNAL Nature 419 (6906), 527-531 (2002)
MEDLINE 22255708
PUBMED 12368867
REFERENCE 3 (bases 1 to 113880)
AUTHORS Lawson,D., Bowman,S. and Barrell,B.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-1998) P.falciparum Genome Sequencing Consortium, The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK
COMMENT On Apr 30, 2000 this sequence version replaced gi:4493924.
For more information about this sequence or the Malaria Project, see http://www.sanger.ac.uk/Projects/P_falciparum.
FEATURES Location/Qualifiers
source 1. .113880
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="3"
/clone="MAL3P4"
gene 2. .1024
/gene="MAL3P4.27"
/note="synonym: PFC0430w"
CDS <2. .1024
/gene="MAL3P4.27"
/note="Revised: removed 5' exon"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD49086.1"
/db_xref="GI:23476984"
/db_xref="SPTREMBL:Q8I222"
/translation="KNYDSIQDLIKSIDFWITLKENIKNIDIPTFQKYLIEKYNKIIK
NEMDVCYVDIVGCIENEEFMKKNEHKCNGYLKSLKEKYEKNNDINFEEECKYINENFI
EEIFTSDYSSPNLDIKNEKDISSDNKNKNNDNNNNNNNNNDDDFHNNNSYIQNND
FLEENIIRKKFIYIPKNDSNNIKIQSNIKQTYKDEKQNKSFLQENINILKSFPQSIQN
LSVEMFYRDNYFYFHCLNKYIKSKNKIEGLIICKLLTYLDLLRSYNTHGLNKTFQRLM
HIEEKIDNNTKGIGETLNTYSCDINPVLANNTSYYGEGPKKYKIIYSIPGKKIKKNKI
KKNKIK"
misc_feature 1218. .1309
/note="Possible exon as peptide fragment matches one of the orphan peptides from proteomics study (source: M. Mann)
maps at the 3' dowmstream region of gene PFC0430w; possibly identifies a small exon II of PFD0430w; could also be a separate very small CDS since it is part of an small open reading frame (ORF)."
gene 2364. .6248
/gene="MAL3P4.26"
/note="synonym: PFC0435w"
CDS 2364. .6248
/gene="MAL3P4.26"
/note="Signal peptide predicted by SignalP 2.0 HMM (Signal peptide probabilty 0.995, signal anchor probability 0.005) with cleavage site probability 0.397 between residues 17 and 18"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAA15616.1"
/db_xref="GI:3764023"
/db_xref="SPTREMBL:O77361"
/translation="MKVGIIFFCLFFFVVLGACNNVKERIFKNIKKRTKFIILNEPIV
DLSFSENLFHTLLFDLDVDKNLYTLDESLLNLENLNYSSIFRLLVDTYKNIKENEDDN
KNIRYIFLGTSFSRIHPLNFEYFLRKLNKYIYNGNIYEKGNVDIRGILEEYNKEIEEK
KLEKQKLNKIKDKNNNNNNNNNSKFSKDGDNEDFNNKNDLYNPSDKLYNNNDDIDVHE
LLEEIITKEKRFFLNDDDDNDSNDKYILKTDEVNKYKGFFIGYGFNDDIPSVIHHYNF
DKNFLFPSLNSGIILDITLLKNIYEVSNILLSNNEKDQSIHIDYIYEVTKYIKENLRV
RLTHSENVCLNEEQNIHLLDNDPNNFEIYKYYQVLNLFKDYNKNTEEKQYEKIGHENV
RHEETSSEGNENLNRNTKHNNDNNNDNNNYSEDAIAELLLSYFNVFYPISTCMCYSIR
SKHESLMDYDKYHMINLENDIKLKHYIKETEEIHFNSIEEYKMKLNRINYKYDTLLEE
HENLVTHKNILIGIKTSINTEEERIPHIKNTYDNKENTQIIFNTFNYDNKLKEKNTFG
FYNNSLLQNALENDNIDLDIIYMSDKESQKYDNLYFNSKVTSKEGLCEKLKHMIYYYY
EEYVMKNSEKKYFFIADDDTFVNVKNLIDVTNLTLNTCSHSKKYMYDKYIKSYDFVKE
NEALFLQNFPKKTLFLYSYLKDTFAKTIQTLKKYDYVPKYCQGGILSKKHKNNDSDDD
HDHHVGNKQNNDSTNHQDIEKNQVNVINNNNNNNNNKAKSIPIYLGRRYSYNTFSTNS
NEYFYDYLTGGAGILINDETAKRIYECKECTCPSTNSSMDDMIFGKWAKELGILAINF
EGYFQNSPLDYNKKYINTLVPITYHRLNKNRTTKESRDMYFNYLVNYNRNDKEQNKDI
YVDYLDRNHKNMIDNVFHYFFYVNMYDEKNKVVTKIEHNADMNSKKNKSKNPQKLNNT
QGDKNVNDDENVNDDENVKGDENVKGDENVKGDEYMKGDENVKGDENVKDDENVKDDE
NIKGDDNNYNVDNMENIDDIINMVESVDDDVMERNKKGTGKEKKDDKNHNNKEKATDV
KKSSVPTNNIDKNEDTTKYVIKMNEKIYNRMQESGKYKQLFDINKFFKKEIEGHPYFQ
KIKKKNEKAKKEKEKMNQLKKQKDYTNNYFHTSNMQGNFNQQKMGNYQNQENEENDFF
DQRPEIEEDAINPMDYEEYMENLSNFEDDGEPYDEYDDYDDFVNTINADKLKINDQNK
HLYEQIKDIAQPPVNFQNDQNSNTFDFDTDEL"
gene complement(7161. .13970)
/gene="MAL3P4.25"
/note="synonym: PFC0440c"
CDS complement(7161. .13970)
/gene="MAL3P4.25"
/note="Similarity to helicases. C.elegans helicase (WP:F52B5.3) BLAST Score: 290, sum P(4) = 5.9e-31; 34% in 199 aa overlap
Pfam match to entry PF00271 helicase_C, Helicase conserved C-terminal domain, score 48.20, E-value 1.8e-10"
/codon_start=1
/product="helicase, putative"
/protein_id="CAA15615.1"
/db_xref="GI:3764022"
/db_xref="GOA:O77360"
/db_xref="SPTREMBL:O77360"
/translation="MNNTFKINKKRRTVYESYNIINIANKKRYMCNDNKMDNDNNIIN
EKKYFNNSSSIKNTGNNYKNDNIDKSYDMCERRSMMDRDENIYNNNYKNKNRNNYSNK
NNYSNNINHMDNRRNTYKGDLDGRKQYLSNKHYESKNDNENSYNLKKYNFNIPPIKYN
NNNNNNNNNNNNDRQNIDNNIYKFYDDKNEKKSNKTHSYGNNFNNEFNTFNDPKQNYM
SNESHNMTQDGHPIYNHNNNNNDGRIKNRKDISLSSYGYNKISNDNLSKENTSYINQY
NNSGGANRHVINAQHNNNQHNNNQHNNNFNYHNYKNAKPYQDKYKNGDSFKFSAYNAE
QHDNMDKNYYQDIHKNNMKHSTKGDTAISFNGQKVFGSFKNKQMKEKRKFFKNKERGS
SINRPGMHNNNNNYNNNNNNNYYNNYNNHFCDDAERRNNKNENWCEINSNELYEKDGD
MPIYNANENCNLKKMIYKNDKKEKEKIKILLTEDNVNNVQSNIEKLSIYKSRNEIIEM
IEKNDVTFINGETGSGKSTCVPKFLLEENIRENKKINIIVTEPRRIACIALSKILSEL
TNEKLGQKIGYRISGESLYDSEKTVITYITIGYLFKLFLHHKNMYKKFTHVIIDEIHD
RSILLDIVLLFIKLYLHNKQKDEQMFKLIIMSATMQSNLFYSYFEHPNIKMGSIFIGT
KIFSIDTFYIEDIINYTRYGSRNIHDEKINDEKINDEKINDEKINDEKINDDNKKSCD
DKKSCDDKKSCDGNNNFVKESVIDFILRRKNCNKINLSKNSEMLLYKIKSEYDKNIHM
FNNNNNKYCNDKNKDLDVDEIIPANVFSNISNLCLELVYNLCLKGDSVIIFLSGMQDI
TDMYHQLSMIINNNMDNPSNSNININNGNDANDANDVNNMHNMHNMHNMHNIANQLTY
HRTDIKIHIHMLHSCLYDNTIHKLKHNDTDINIFLSSNIAESSITIPNVRLVIDFCIQ
KNIEYNDKKKAHILVKKWINKSSMEQRKGRCGRTCHGICIRMISKNFLNLLRDHKISE
IYTHSLHLLYLYILKSMSVLNGLINKRNEVVHCDNTNVYNVKEVKDQKNGICKNEISH
SHAENKKLTIYDVLSMIIEKPSKEKIKSTRYELEKVKAVIKIKDKLVISIIGQIMIRF

NLSINLCRLLLYGVLLDVTFDTIILIGILNTNDIFPNINLYSSKNIYSYAVSLEICSK
QKAYFDGNTYSEPIMLRNVFLEWLCVYLLYVQGLKKENKFHRKELKTYYMNTCSIMNK
RNHVNSKKLLCVINSVDNLCKKMLKMLNKNSNAYKSCVYLLYLLRGGSDINYNMGNNS
NNNNDNSGEIINADGITPNNTHMNEINGMNVFNVVTKYCLFDYSNQNIYLKFLFSLSF
TPLFIHGSPNLPLRDLNEGKKKSKKLMNVLNFMIDNKLDVKNCIYFSGVYIPDINILK
RAICIMCPYLSFDIYVSKNIYYVYFHNNNYHNNYHNNILKGDDFSYYNKTICDKGDAD
MSVNSYLNLTQEGHDTDGITSFGVSPLMNNNNMYGNNYIHNNVYYPCNGINNYYNNNV
VNSQNNYMMKMIMSNYYSESFLNLTSGKNYPDIIEKIRNKLLYKYNNNVEKSIFHNFY
NNILNLIQPVNAIVPVYSNKYDEINNASICTNIINMFSNGKWIFTMPLFNPNKRKCEV
SEYFNHKYELKKPKHLFLVKWMFLNTDNYKIKKNKQKEHVMDTNNALVNNNSLGKNVS
KNDIHRDNINDYSLNDQEDQQYNKGYYYDDYDDYDDYDDDDNNDDNNDNNNNNNDDNN
NNNDDDYLSNVKVESDNNNTHQNNPDGYEGNDSSSMCSDNNKKSKQKKMKCNLNFRSV
LGFLSMCPFNFDTKRNIYDNQTADIYAVCSSIDYSYTNDTYTWVNYVTVLPNKYFLTF
FLSSIPYHDNVIIQTRTNLYTSDILSIKIFDSKEIILKNMKKMKHKNKTNTTNYYDTN
STINNPNITKHDLLRVNYLRYVMSKLLLSLTLSFNRQEIEQNTQPLKLSEKSKRKSVI
INKDGPNNDNRVDANDKNVNGNDKNVNDNDKNVNSQDNQGNQKDHFDDEDEEYFSDED
IMSDIEENNILKENLYLKQIVINMIEENNMDYINYNKYSMDNMDIDEHFNLQYLYEDE
HFINLNEKWYNNKGEILMNTYDYLYNSLKNENDYNQYNNNNNNNNNNNNNSSYMVKNI
KEKALNSLEAVKSSYLNNNFSVQVQNTMNAYERTNSEDFLFFQPINISPIKEANYKLR
SMYYNEVAPTLDKFRRTNRQV"

variation 15740
/note="variable sequence, 9 or 10 Ts"

gene complement(join(16039. .16667,16813. .16891))
/gene="PFC0441c"

CDS complement(join(16039. .16667,16813. .16891))
/gene="PFC0441c"
/note="Revised: new gene prediction; splicing confirmed in P. yoelii"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB89483.1"
/db_xref="GI:7672213"
/db_xref="SPTREMBL:Q9NFE6"
/translation="MDINNIMEEVQAIKVLPHMMFNASKLKVLSDYVNVAFENNEYYD
NDVMLTLLRLFCLYPHCYDKDIIKKILICVLYNINNVDMNMYMSLINSSLYDDNIKSV
IYIYELIKNCQYKKLWNCINNNIGNEYNNCDYSYLKNDKNFICNIRKYILNTISISFE
SIYLKNMSEYLNIQDNTQLEQFLNENKWTIKMINHKGKDEQICYNGNIETVQNKKNIN
TYFTEDNIGSYITKLNH"

gene join(18192. .18242,18754. .19140)
/gene="MAL3P4.24"
/note="synonym: PFC0445w"

Query Match 26.7%; Score 28.6; DB 3; Length 113880;
Best Local Similarity 21.8%; Pred. No. 3.2e+02;
Matches 22; Conservative 12; Mismatches 67; Indels 0; Gaps 0;

```
Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              ||||:      |:|:            :::      ||||| :| ||| |:      ::
Db      51946 AAAATTATAAATATATTTTGTTAACTTATTTTATTATAAAAATTATAAATATATTTTGTT 51887

Qy         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUA 103
               |          | |||:          : |
Db      51886 AACATTTTTTTTTTATAAAAATTATAAATATATTTTGTTAA 51846
```

RESULT 15
AC099283/c
LOCUS AC099283 227182 bp DNA linear HTG 13-MAY-2003
DEFINITION Rattus norvegicus clone CH230-133B5, WORKING DRAFT SEQUENCE, 3 unordered pieces.
ACCESSION AC099283
VERSION AC099283.6 GI:30579809
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 227182)
AUTHORS Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwackelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 227182)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (09-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 227182)
AUTHORS Rat Genome Sequencing Consortium.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On May 13, 2003 this sequence version replaced gi:25086297.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
---------- Genome Center

```
            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
          -------------- Project Information
            Center project name: GKCS
            Center clone name: CH230-133B5
          -------------- Summary Statistics
            Assembly program: Atlas 3.0;
            Consensus quality: 221120 bases at least Q40
            Consensus quality: 222542 bases at least Q30
            Consensus quality: 223333 bases at least Q20
            Estimated insert size: 228594; sum-of-contigs estimation
            Quality coverage: 9x in Q20 bases; sum-of-contigs estimation
          --------------
          * NOTE: Estimated insert size may differ from sequence length
          *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 3 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *       1   224728: contig of 224728 bp in length
          *  224729   224828: gap of unknown length
          *  224829   225852: contig of 1024 bp in length
          *  225853   225952: gap of unknown length
          *  225953   227182: contig of 1230 bp in length.
'EATURES             Location/Qualifiers
     source          1. .227182
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-133B5"
     misc_feature    1. .1303
                     /note="wgs_end_extension
                     clone_end:Sp6"
     misc_feature    2184. .3927
                     /note="wgs_end_extension
                     clone_end:Sp6"
     misc_feature    5122. .5977
                     /note="clone_boundary
                     clone_end:Sp6
                     site:EcoRI
                     end_sequence:BH323630"
     misc_feature    221329. .222127
                     /note="clone_boundary
                     clone_end:T7
                     site:EcoRI
                     end_sequence:BH323591"
)RIGIN

 Query Match             26.7%;  Score 28.6;  DB 2;  Length 227182;
 Best Local Similarity   21.8%;  Pred. No. 2.6e+02;
 Matches   22;  Conservative   12;  Mismatches   67;  Indels    0;  Gaps    0;

)y          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              ||  :     : |:|          :   :::      ||||| :| ||| |:     |:
)b     109443 AATGTTATTATTATAAAATTCTGATGTTTTTTGAAGAAAAAAATAGAAAAATCTATCCTA 109384

)y         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUA 103
               ||           | |||                 ::|
)b     109383 TAGACAGTCATAGAAAATAAAGGACTCAAAAAAAACTCTTA 109343


earch completed: March  8, 2004, 05:20:24
ob time : 2071 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

)M nucleic - nucleic search, using sw model

Run on: March 8, 2004, 03:12:12 ; Search time 307 Seconds
(without alignments)
1480.642 Million cell updates/sec

Title: US-09-646-651C-2
Perfect score: 107
Sequence: 1 ggaaaaunnnnnunauaugn.....nnnnnnnnnnnnnuuagcag 107

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	31.4	29.3	298	2	AAX23773	Aax23773 Ribonucle
c	2	29.4	27.5	110000	2	AAV21209_05	Continuation (6 of
	3	29.2	27.3	3456	4	ABL18738	Abl18738 Drosophil
	4	29.2	27.3	4829	4	ABL18730	Abl18730 Drosophil
	5	27.4	25.6	524	6	ABQ89024	Abq89024 Human pro
	6	27.4	25.6	524	8	ADB81968	Adb81968 Human cDN
	7	27.4	25.6	761	8	ADB82181	Adb82181 Human cDN
	8	27.4	25.6	769	6	ABQ89163	Abq89163 Human pro
	9	27.4	25.6	769	8	ADB82106	Adb82106 Human cDN
	10	27.4	25.6	789	6	ABQ89160	Abq89160 Human pro
	11	27.4	25.6	789	8	ADB82103	Adb82103 Human cDN
	12	27.4	25.6	835	6	ABQ89238	Abq89238 Human pro
	13	27.4	25.6	929	8	ADB83094	Adb83094 Human cDN
	14	27	25.2	325791	4	AAS43104	Aas43104 Human Oes
c	15	26.8	25.0	6419	6	ABL32266	Abl32266 Human imm
c	16	26.8	25.0	13814	6	ABL33193	Abl33193 Human imm
	17	26.6	24.9	10034	6	ABL33885	Abl33885 Human imm
	18	26.2	24.5	100543	6	ABS52816	Abs52816 Genomic D
	19	26	24.3	318	8	ADA29355	Ada29355 DNA encod
c	20	26	24.3	5504	6	ABL70350	Abl70350 Chemicall
c	21	26	24.3	5504	6	AAS61309	Aas61309 Human gen
	22	26	24.3	7977	6	ABL34089	Abl34089 Human imm
	23	25.8	24.1	250	3	AAC16119	Aac16119 Human sec
	24	25.8	24.1	553	7	ABX65321	Abx65321 Human gen
c	25	25.8	24.1	934	2	AAT92699	Aat92699 Candida C
c	26	25.8	24.1	934	2	AAT92866	Aat92866 Candida C
c	27	25.8	24.1	934	2	AAZ11688	Aaz11688 Candida R
c	28	25.8	24.1	1386	7	ACA27482	Aca27482 Prokaryot
	29	25.8	24.1	1905	7	ACA29255	Aca29255 Prokaryot
c	30	25.8	24.1	3243	2	AAV32555	Aav32555 Candida a
	31	25.8	24.1	4544	6	AAD48259	Aad48259 Ehrlichia
c	32	25.8	24.1	7461	4	AAS30639	Aas30639 DNA encod
c	33	25.8	24.1	7461	4	AAS28701	Aas28701 Genomic s
c	34	25.8	24.1	7461	7	ACA03402	Aca03402 DNA encod
c	35	25.8	24.1	7461	8	ADB96750	Adb96750 Novel lun
	36	25.8	24.1	10205	6	ABK31275	Abk31275 Signal tr
	37	25.8	24.1	10205	6	ABL70236	Abl70236 Chemicall
	38	25.8	24.1	12639	6	ABN80106	Abn80106 Human che
	39	25.8	24.1	12781	4	AAS46386	Aas46386 Tumour su
	40	25.8	24.1	12781	6	ABL92229	Abl92229 Chemicall
	41	25.8	24.1	12781	9	ADB54256	Adb54256 Pretreate
	42	25.8	24.1	12781	9	ADB54128	Adb54128 Pretreate
c	43	25.8	24.1	32174	4	AAS30638	Aas30638 DNA encod
c	44	25.8	24.1	32174	4	AAI62606	Aai62606 Human bre
c	45	25.8	24.1	32174	4	AAL36280	Aal36280 Human mus

ALIGNMENTS

RESULT 1
AAX23773
ID AAX23773 standard; RNA; 298 BP.
XX
AC AAX23773;
XX
DT 21-JUN-1999 (first entry)
XX
DE Ribonucleotide protein ARNA VI Clone P10 RNA.
XX
KW Angiotropin related protein; ARP; ternary complex; S100 protein;
KW copper-containing ribonucleoprotein; copper; cell selective;
KW morphogenic action; blood capillary endothelial cell; confluent;
KW non-mitogenic induction; cell phenotype; three-dimensional organoid;
KW spatiotemporal supracellular organisation; chemotropic; blood vessel;
KW tissue neovascularisation; angiogenesis modulation; ss.
XX
OS Synthetic.
XX
PN DE19811047-C1.
XX
PD 15-APR-1999.
XX
PF 13-MAR-1998; 98DE-01011047.
XX
PR 13-MAR-1998; 98DE-01011047.
XX
PA (FRAU) FRAUNHOFER GES FOERDERUNG ANGEWANDTEN.
XX
PI Kiesewetter S, Kuhn E, Koch-Pelster B, Brunner H;
XX
DR WPI; 1999-216114/19.
XX
PT Copper-containing ribonucleoproteins - useful for modulating
PT angiogenesis.
XX
PS Claim 2; Page 12; 16pp; German.
XX
CC This invention describes novel copper-containing ribonucleoproteins which
CC are ternary complexes of an S100 protein, copper ions, and RNA comprising
CC the following consensus sequence or its complement GGAAAAUNNNNNUNAUAUGN1-
CC 6CUNNNUUUNNNNNNAAAAN0-1UANAAACAUN0-5CUUNAGN0- 13AGAA-AUN0-16UUAGCAG where
CC N = G, A, U or C. The ribonucleoproteins are stated to have the following
CC properties (1) cell-selective morphogenic action in vitro on isolated
CC primary and/or cloned blood capillary endothelial cells in culture for
CC the non-mitogenic induction of the change in cell phenotype from the

```
CC   confluent state, for non-mitogenic alteration of the spatiotemporal
CC   supracellular organisation of cells into three-dimensional organoid,
CC   capillary-like structures in culture, (2) a specific chemotropic action
CC   on blood vessels in vivo, (3) induction of directional growth of blood
CC   vessels in vivo and (4) induction of neovascularisation of tissues
CC   through directed ingrowth of blood vessels. Their use for modulating
CC   angiogenesis is claimed
XX
SQ   Sequence 298 BP; 103 A; 40 C; 38 G; 0 T; 117 U; 0 Other;

  Query Match             29.3%;  Score 31.4;  DB 2;  Length 298;
  Best Local Similarity   42.1%;  Pred. No. 6.2;
  Matches   32;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy         32 UUNNNNNNAAAAANUANAAACAUNNNNNCUUNAGNNNNNNNNNNNNNNAGAAAUNNNNNNN 91
              ||        |||| || ||||||     ||| ||              ||||||
Db         89 UUUGUACAUAAAAAUAAAAACAUGACUUCUUUAGACACUCCUUCAUUAGAAAUAAAAUAA 148

Qy         92 NNNNNNNNNUUAGCAG 107
                      |||||||
Db        149 AAUAAACUAUUAGCAG 164


RESULT 2
AAV21209_05/c
Continuation (6 of 17) of AAV21209 from base 500001 (Methanococcus jannaschii circular c
WP Sequence split into 17 fragments  LOCUS AAV21209 Accession Aav21209
WP      Fragment Name       Begin     End
WP      AAV21209_00             1    110000
WP      AAV21209_01        100001    210000
WP      AAV21209_02        200001    310000
WP      AAV21209_03        300001    410000
WP      AAV21209_04        400001    510000
WP      AAV21209_05        500001    610000
WP      AAV21209_06        600001    710000
WP      AAV21209_07        700001    810000
WP      AAV21209_08        800001    910000
WP      AAV21209_09        900001   1010000
WP      AAV21209_10       1000001   1110000
WP      AAV21209_11       1100001   1210000
WP      AAV21209_12       1200001   1310000
WP      AAV21209_13       1300001   1410000
WP      AAV21209_14       1400001   1510000
WP      AAV21209_15       1500001   1610000
WP      AAV21209_16       1600001   1664976

  Query Match             27.5%;  Score 29.4;  DB 2;  Length 110000;
  Best Local Similarity   24.8%;  Pred. No. 47;
  Matches   26;  Conservative   10;  Mismatches   69;  Indels    0;  Gaps    0;

Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              ||||:     : | | |      |:   :::      |||||:|| |||||:     :
Db      73299 AAAATTAGTATTAGAAGCTATAACTTATTTTAATGCCAAAAAATATAAACATTTCTGGTG 73240

Qy         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAGCAG 107
               |              ||  |:                  | ||
Db      73239 AAATAATGACAGGCAAAGTTATCTTAGTAGGAGCAGGACCGGGAG 73195


RESULT 3
ABL18738
ID   ABL18738 standard; DNA; 3456 BP.
XX
AC   ABL18738;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster genomic polynucleotide SEQ ID NO 7687.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ds.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US009231.
XX
PR   23-MAR-2000; 2000US-0191637P.
PR   11-JUL-2000; 2000US-00614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signaling and cell-cell
PT   interactions.
XX
PS   Claim 1; SEQ ID NO 7687; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC   ABB72072). The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 3456 BP; 827 A; 939 C; 888 G; 802 T; 0 U; 0 Other;

  Query Match             27.3%;  Score 29.2;  DB 4;  Length 3456;
  Best Local Similarity   25.0%;  Pred. No. 34;
  Matches   21;  Conservative   13;  Mismatches   50;  Indels    0;  Gaps    0;

Qy          1 GGAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNC 60
              ||||||:     :  :|:       |:   :::      |||||:| |||  :
Db        769 GGAAAATAAACTTAGTATTTACAATCTTAGTTTTGTACCAAAAAATAAAAATCTGTATAA 828

Qy         61 UUNAGNNNNNNNNNNNNNAGAAAU 84
              :: |                |||:
Db        829 TTAATTTTAAATTCAAATTTAAAT 852


RESULT 4
ABL18730
ID   ABL18730 standard; DNA; 4829 BP.
XX
AC   ABL18730;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster genomic polynucleotide SEQ ID NO 7663.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ds.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US009231.
XX
PR   23-MAR-2000; 2000US-0191637P.
PR   11-JUL-2000; 2000US-00614150.
XX
PA   (PEKE ) PE CORP NY.
```

```
XX
PI   Venter JC, Adams M, Li PWD, Myers EW;
XX
DR   WPI; 2001-656860/75.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signaling and cell-cell
PT   interactions.
XX
PS   Claim 1; SEQ ID NO 7663; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC   ABB72072). The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 4829 BP; 1251 A; 1183 C; 1154 G; 1241 T; 0 U; 0 Other;

  Query Match             27.3%;  Score 29.2;  DB 4;  Length 4829;
  Best Local Similarity   25.0%;  Pred. No. 36;
  Matches   21;  Conservative   13;  Mismatches   50;  Indels    0;  Gaps    0;

Qy          1 GGAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNC 60
              ||||||:     :  :|:      |:   :::      |||||:|||  :
Db       3811 GGAAAATAAACTTAGTATTTACAATCTTAGTTTTGTACCAAAAAATAAAAATCTGTATAA 3870

Qy         61 UUNAGNNNNNNNNNNNNNAGAAAU 84
              :: |              |||:
Db       3871 TTAATTTTAAATTCAAATTTAAAT 3894


RESULT 5
ABQ89024
ID   ABQ89024 standard; cDNA; 524 BP.
XX
AC   ABQ89024;
XX
DT   27-SEP-2002  (first entry)
XX
DE   Human prostate expressed polynucleotide SEQ ID NO 280.
XX
KW   Human; prostate; cytostatic; tumour; cancer; vaccine; gene therapy; gene;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200255700-A2.
XX
PD   18-JUL-2002.
XX
PF   07-DEC-2001; 2001WO-US047349.
XX
PR   07-DEC-2000; 2000US-0254648P.
PR   13-MAR-2001; 2001US-0275688P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (HYSE-) HYSEQ INC.
XX
PI   Escobedo J, Garcia PD, Kassam A, Lamson G, Drmanac R;
PI   Crkvenjakov R, Dickson M, Drmanac S, Labat I, Leshkowitz D, Kita D;
PI   Garcia V, Jones WL, Stache-Crain B, Scott EM;
XX
DR   WPI; 2002-557824/59.
XX
PT   New genes and gene products isolated from human prostate, useful for
PT   treating or diagnosing tumor or cancer (e.g. prostate cancer or breast
PT   cancer), or as vaccines for treating or preventing these diseases.
XX
PS   Claim 1; SEQ ID NO 280; 186pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated polynucleotide comprising any of
CC   1477 sequences or its fragment, degenerate variant, antisense or
CC   complement. The polynucleotides and gene products are useful for treating
CC   or diagnosing tumour or cancer (e.g. prostate cancer, breast cancer, lung
CC   cancer or medullary carcinoma) in a subject (e.g. cattle, dogs, cats,
CC   rabbits, horse or human). The polynucleotides and polypeptides are also
CC   useful as vaccines for treating or preventing these diseases. The
CC   polynucleotides are useful for gene therapy. The present sequence is that
CC   of one of a group of polynucleotides (ABQ88745-ABQ90015) disclosed
CC   electronically as sequences of the invention. However only 1271
CC   polynucleotide sequences are given, whereas 1477 polynucleotides and 91
CC   proteins are claimed. Note: The sequence data for this patent did not
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequence
XX
SQ   Sequence 524 BP; 132 A; 107 C; 110 G; 159 T; 0 U; 16 Other;

  Query Match             25.6%;  Score 27.4;  DB 6;  Length 524;
  Best Local Similarity   24.3%;  Pred. No. 87;
  Matches   25;  Conservative    9;  Mismatches   69;  Indels    0;  Gaps    0;

Qy          2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCU 61
              |||||      : | |        :    :       ||||  :| |||| :     |:
Db        232 GAAAACTCTTGTTAGACATGATGTTTCCTATACAGAAGAAAAGATACAAACTTGATTGCT 291

Qy         62 UNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUAG 104
              : ||               |||  |:              : ||
Db        292 TCAGTAAAAACACTTTTAGAGATTGTGTTTCCTTTTTTGTGAG 334


RESULT 6
ADB81968
ID   ADB81968 standard; cDNA; 524 BP.
XX
AC   ADB81968;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Human cDNA sequence useful for the treatment of cancer (SeqID 280).
XX
KW   human; prostate; cancer; cytostatic; gene therapy; vaccine;
KW   immune response; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2003050236-A2.
XX
PD   19-JUN-2003.
XX
PF   04-SEP-2002; 2002WO-US028214.
XX
PR   07-DEC-2001; 2001US-00012697.
XX
PA   (CHIR ) CHIRON CORP.
PA   (HYSE-) HYSEQ INC.
XX
PI   Escobedo J, Garcia PD, Kassam A, Lamson G, Drmanac R;
PI   Crkvenjakov R, Dickson M, Drmanac S, Labat I, Leshkowitz D, Kita D;
PI   Garcia V, Jones LW, Stache-Crain B, Scott EM;
XX
DR   WPI; 2003-513972/48.
XX
PT   New polynucleotides derived from human prostate, useful for modulating
PT   immune response to prevent or treat cancer.
XX
PS   Claim 1; SEQ ID NO 280; 188pp; English.
XX
CC   This invention relates to novel isolated polynucleotides of human origin,
CC   particularly isolated from the human prostate. Specifically, it refers to
CC   the diagnostics and therapeutics comprising these novel human
```

CC polynucleotides, and includes the derived probes, antisense
CC oligonucleotides and antibodies thereof. The identification of these
CC human prostate genes that can inhibit tumour growth is useful for
CC understanding the progression and nature of complex diseases such as
CC cancer, and hence they are important in the drug discovery process. The
CC present invention describes these polynucleotides and encoded
CC polypeptides as exhibiting cytostatic activity, and through gene therapy
CC and/ or vaccines they can be used to modulate the immune response for the
CC prevention or treatment of cancers, particularly of the prostate, but
CC also for breast, lung and colon cancer. This polynucleotide sequence is a
CC human cDNA sequence useful for the treatment of cancer, used in an
CC exemplification of the invention. NOTE: These sequences are not given in
CC the specification but are provided on the WIPO website.
XX
SQ Sequence 524 BP; 132 A; 107 C; 110 G; 159 T; 0 U; 16 Other;

Query Match 25.6%; Score 27.4; DB 8; Length 524;
Best Local Similarity 24.3%; Pred. No. 87;
Matches 25; Conservative 9; Mismatches 69; Indels 0; Gaps 0;

```
Qy          2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCU 61
              |||||      : | |         :    :       ||||  :| |||| :     |:
Db        232 GAAAACTCTTGTTAGACATGATGTTTCCTATACAGAAGAAAAGATACAAACTTGATTGCT 291

Qy         62 UNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAG 104
              : ||              ||| |:              : ||
Db        292 TCAGTAAAAACACTTTTAGAGATTGTGTTTCCTTTTTTGTGAG 334
```

RESULT 7
ADB82181
ID ADB82181 standard; cDNA; 761 BP.
XX
AC ADB82181;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human cDNA sequence useful for the treatment of cancer (SeqID 493).
XX
KW human; prostate; cancer; cytostatic; gene therapy; vaccine;
KW immune response; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003050236-A2.
XX
PD 19-JUN-2003.
XX
PF 04-SEP-2002; 2002WO-US028214.
XX
PR 07-DEC-2001; 2001US-00012697.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Escobedo J, Garcia PD, Kassam A, Lamson G, Drmanac R;
PI Crkvenjakov R, Dickson M, Drmanac S, Labat I, Leshkowitz D, Kita D;
PI Garcia V, Jones LW, Stache-Crain B, Scott EM;
XX
DR WPI; 2003-513972/48.
XX
PT New polynucleotides derived from human prostate, useful for modulating
PT immune response to prevent or treat cancer.
XX
PS Claim 1; SEQ ID NO 493; 188pp; English.
XX
CC This invention relates to novel isolated polynucleotides of human origin,
CC particularly isolated from the human prostate. Specifically, it refers to
CC the diagnostics and therapeutics comprising these novel human
CC polynucleotides, and includes the derived probes, antisense
CC oligonucleotides and antibodies thereof. The identification of these
CC human prostate genes that can inhibit tumour growth is useful for
CC understanding the progression and nature of complex diseases such as
CC cancer, and hence they are important in the drug discovery process. The
CC present invention describes these polynucleotides and encoded
CC polypeptides as exhibiting cytostatic activity, and through gene therapy
CC and/ or vaccines they can be used to modulate the immune response for the
CC prevention or treatment of cancers, particularly of the prostate, but
CC also for breast, lung and colon cancer. This polynucleotide sequence is a
CC human cDNA sequence useful for the treatment of cancer, used in an
CC exemplification of the invention. NOTB: These sequences are not given in
CC the specification but are provided on the WIPO website.
XX
SQ Sequence 761 BP; 197 A; 167 C; 168 G; 214 T; 0 U; 15 Other;

Query Match 25.6%; Score 27.4; DB 8; Length 761;
Best Local Similarity 24.3%; Pred. No. 91;
Matches 25; Conservative 9; Mismatches 69; Indels 0; Gaps 0;

```
Qy          2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCU 61
              |||||      : | |         :    :       ||||  :| |||| :     |:
Db        237 GAAAACTCTTGTTAGACATGATGTTTCCTATACAGAAGAAAAGATACAAACTTGATTGCT 296

Qy         62 UNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAG 104
              : ||              ||| |:              : ||
Db        297 TCAGTAAAAACACTTTTAGAGATTGTGTTTCCTTTTTTGTGAG 339
```

RESULT 8
ABQ89163
ID ABQ89163 standard; cDNA; 769 BP.
XX
AC ABQ89163;
XX
DT 27-SEP-2002 (first entry)
XX
DE Human prostate expressed polynucleotide SEQ ID NO 419.
XX
KW Human; prostate; cytostatic; tumour; cancer; vaccine; gene therapy; gene;
KW ss.
XX
OS Homo sapiens.
XX
PN WO200255700-A2.
XX
PD 18-JUL-2002.
XX
PF 07-DEC-2001; 2001WO-US047349.
XX
PR 07-DEC-2000; 2000US-0254648P.
PR 13-MAR-2001; 2001US-0275688P.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Escobedo J, Garcia PD, Kassam A, Lamson G, Drmanac R;
PI Crkvenjakov R, Dickson M, Drmanac S, Labat I, Leshkowitz D, Kita D;
PI Garcia V, Jones WL, Stache-Crain B, Scott EM;
XX
DR WPI; 2002-557824/59.
XX
PT New genes and gene products isolated from human prostate, useful for
PT treating or diagnosing tumor or cancer (e.g. prostate cancer or breast
PT cancer), or as vaccines for treating or preventing these diseases.
XX
PS Claim 1; SEQ ID NO 419; 186pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide comprising any of
CC 1477 sequences or its fragment, degenerate variant, antisense or
CC complement. The polynucleotides and gene products are useful for treating
CC or diagnosing tumour or cancer (e.g. prostate cancer, breast cancer, lung
CC cancer or medullary carcinoma) in a subject (e.g. cattle, dogs, cats,
CC rabbits, horse or human). The polynucleotides and polypeptides are also
CC useful as vaccines for treating or preventing these diseases. The
CC polynucleotides are useful for gene therapy. The present sequence is that
CC of one of a group of polynucleotides (ABQ88745-ABQ90015) disclosed

CC electronically as sequences of the invention. However only 1271
CC polynucleotide sequences are given, whereas 1477 polynucleotides and 91
CC proteins are claimed. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequence
XX
SQ Sequence 769 BP; 179 A; 185 C; 196 G; 169 T; 0 U; 40 Other;

Query Match 25.6%; Score 27.4; DB 6; Length 769;
Best Local Similarity 24.3%; Pred. No. 91;
Matches 25; Conservative 9; Mismatches 69; Indels 0; Gaps 0;

Qy 2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCU 61
Db 88 GAAAACTCTTGTTAGACATGATGTTTCCTATACAGAAGAAAAGATACAAACTTGATTGCT 147

Qy 62 UNAGNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUAG 104
Db 148 TCAGTAAAAACACTTTTAGAGATTGTGTTTCCTTTTTTGTGAG 190

RESULT 9
ADB82106
ID ADB82106 standard; cDNA; 769 BP.
XX
AC ADB82106;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human cDNA sequence useful for the treatment of cancer (SeqID 418).
XX
KW human; prostate; cancer; cytostatic; gene therapy; vaccine;
KW immune response; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003050236-A2.
XX
PD 19-JUN-2003.
XX
PF 04-SEP-2002; 2002WO-US028214.
XX
PR 07-DEC-2001; 2001US-00012697.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Escobedo J, Garcia PD, Kassam A, Lamson G, Drmanac R;
PI Crkvenjakov R, Dickson M, Drmanac S, Labat I, Leshkowitz D, Kita D;
PI Garcia V, Jones LW, Stache-Crain B, Scott EM;
XX
DR WPI; 2003-513972/48.
XX
PT New polynucleotides derived from human prostate, useful for modulating
PT immune response to prevent or treat cancer.
XX
PS Claim 1; SEQ ID NO 418; 188pp; English.
XX
CC This invention relates to novel isolated polynucleotides of human origin,
CC particularly isolated from the human prostate. Specifically, it refers to
CC the diagnostics and therapeutics comprising these novel human
CC polynucleotides, and includes the derived probes, antisense
CC oligonucleotides and antibodies thereof. The identification of these
CC human prostate genes that can inhibit tumour growth is useful for
CC understanding the progression and nature of complex diseases such as
CC cancer, and hence they are important in the drug discovery process. The
CC present invention describes these polynucleotides and encoded
CC polypeptides as exhibiting cytostatic activity, and through gene therapy
CC and/ or vaccines they can be used to modulate the immune response for the
CC prevention or treatment of cancers, particularly of the prostate, but
CC also for breast, lung and colon cancer. This polynucleotide sequence is a
CC human cDNA sequence useful for the treatment of cancer, used in an
CC exemplification of the invention. NOTE: These sequences are not given in
CC the specification but are provided on the WIPO website.
XX
SQ Sequence 769 BP; 179 A; 185 C; 196 G; 169 T; 0 U; 40 Other;

Query Match 25.6%; Score 27.4; DB 8; Length 769;
Best Local Similarity 24.3%; Pred. No. 91;
Matches 25; Conservative 9; Mismatches 69; Indels 0; Gaps 0;

Qy 2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCU 61
Db 88 GAAAACTCTTGTTAGACATGATGTTTCCTATACAGAAGAAAAGATACAAACTTGATTGCT 147

Qy 62 UNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUAG 104
Db 148 TCAGTAAAAACACTTTTAGAGATTGTGTTTCCTTTTTTGTGAG 190

RESULT 10
ABQ89160
ID ABQ89160 standard; cDNA; 789 BP.
XX
AC ABQ89160;
XX
DT 27-SEP-2002 (first entry)
XX
DE Human prostate expressed polynucleotide SEQ ID NO 416.
XX
KW Human; prostate; cytostatic; tumour; cancer; vaccine; gene therapy; gene;
KW ss.
XX
OS Homo sapiens.
XX
PN WO200255700-A2.
XX
PD 18-JUL-2002.
XX
PF 07-DEC-2001; 2001WO-US047349.
XX
PR 07-DEC-2000; 2000US-0254648P.
PR 13-MAR-2001; 2001US-0275688P.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Escobedo J, Garcia PD, Kassam A, Lamson G, Drmanac R;
PI Crkvenjakov R, Dickson M, Drmanac S, Labat I, Leshkowitz D, Kita D;
PI Garcia V, Jones WL, Stache-Crain B, Scott EM;
XX
DR WPI; 2002-557824/59.
XX
PT New genes and gene products isolated from human prostate, useful for
PT treating or diagnosing tumor or cancer (e.g. prostate cancer or breast
PT cancer), or as vaccines for treating or preventing these diseases.
XX
PS Claim 1; SEQ ID NO 416; 186pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide comprising any of
CC 1477 sequences or its fragment, degenerate variant, antisense or
CC complement. The polynucleotides and gene products are useful for treating
CC or diagnosing tumour or cancer (e.g. prostate cancer, breast cancer, lung
CC cancer or medullary carcinoma) in a subject (e.g. cattle, dogs, cats,
CC rabbits, horse or human). The polynucleotides and polypeptides are also
CC useful as vaccines for treating or preventing these diseases. The
CC polynucleotides are useful for gene therapy. The present sequence is that
CC of one of a group of polynucleotides (ABQ88745-ABQ90015) disclosed
CC electronically as sequences of the invention. However only 1271
CC polynucleotide sequences are given, whereas 1477 polynucleotides and 91
CC proteins are claimed. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequence
XX
SQ Sequence 789 BP; 199 A; 185 C; 186 G; 181 T; 0 U; 38 Other;

```
  Query Match           25.6%;  Score 27.4;  DB 6;  Length 789;
  Best Local Similarity 24.3%;  Pred. No. 91;
  Matches   25;  Conservative    9;  Mismatches   69;  Indels    0;  Gaps    0;

Qy          2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCU 61
              |||||      : | |         :    :       ||||  :| |||| :     |:
Db         97 GAAAACTCTTGTTAGACATGATGTTTCCTATACAGAAGAAAAGATACAAACTTGATTGCT 156

Qy         62 UNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAG 104
              : ||              ||| |:                : ||
Db        157 TCAGTAAAAACACTTTTAGAGATTGTGTTTCCTTTTTTGTGAG 199


RESULT 11
ADB82103
ID   ADB82103 standard; cDNA; 789 BP.
XX
AC   ADB82103;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Human cDNA sequence useful for the treatment of cancer (SeqID 415).
XX
KW   human; prostate; cancer; cytostatic; gene therapy; vaccine;
KW   immune response; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2003050236-A2.
XX
PD   19-JUN-2003.
XX
PF   04-SEP-2002; 2002WO-US028214.
XX
PR   07-DEC-2001; 2001US-00012697.
XX
PA   (CHIR ) CHIRON CORP.
PA   (HYSE-) HYSEQ INC.
XX
PI   Escobedo J,  Garcia PD,  Kassam A,  Lamson G,  Drmanac R;
PI   Crkvenjakov R,  Dickson M,  Drmanac S,  Labat I,  Leshkowitz D,  Kita D;
PI   Garcia V,  Jones LW,  Stache-Crain B,  Scott EM;
XX
DR   WPI; 2003-513972/48.
XX
PT   New polynucleotides derived from human prostate, useful for modulating
PT   immune response to prevent or treat cancer.
XX
PS   Claim 1; SEQ ID NO 415; 188pp; English.
XX
CC   This invention relates to novel isolated polynucleotides of human origin,
CC   particularly isolated from the human prostate. Specifically, it refers to
CC   the diagnostics and therapeutics comprising these novel human
CC   polynucleotides, and includes the derived probes, antisense
CC   oligonucleotides and antibodies thereof. The identification of these
CC   human prostate genes that can inhibit tumour growth is useful for
CC   understanding the progression and nature of complex diseases such as
CC   cancer, and hence they are important in the drug discovery process. The
CC   present invention describes these polynucleotides and encoded
CC   polypeptides as exhibiting cytostatic activity, and through gene therapy
CC   and/ or vaccines they can be used to modulate the immune response for the
CC   prevention or treatment of cancers, particularly of the prostate, but
CC   also for breast, lung and colon cancer. This polynucleotide sequence is a
CC   human cDNA sequence useful for the treatment of cancer, used in an
CC   exemplification of the invention. NOTE: These sequences are not given in
CC   the specification but are provided on the WIPO website.
XX
SQ   Sequence 789 BP; 199 A; 185 C; 186 G; 181 T; 0 U; 38 Other;

  Query Match           25.6%;  Score 27.4;  DB 8;  Length 789;
  Best Local Similarity 24.3%;  Pred. No. 91;
  Matches   25;  Conservative    9;  Mismatches   69;  Indels    0;  Gaps    0;

Qy          2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCU 61
              |||||      : | |         :    :       ||||  :| |||| :     |:
Db         97 GAAAACTCTTGTTAGACATGATGTTTCCTATACAGAAGAAAAGATACAAACTTGATTGCT 156

Qy         62 UNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAG 104
              : ||              ||| |:                : ||
Db        157 TCAGTAAAAACACTTTTAGAGATTGTGTTTCCTTTTTTGTGAG 199


RESULT 12
ABQ89238
ID   ABQ89238 standard; cDNA; 835 BP.
XX
AC   ABQ89238;
XX
DT   27-SEP-2002  (first entry)
XX
DE   Human prostate expressed polynucleotide SEQ ID NO 494.
XX
KW   Human; prostate; cytostatic; tumour; cancer; vaccine; gene therapy; gene;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200255700-A2.
XX
PD   18-JUL-2002.
XX
PF   07-DEC-2001; 2001WO-US047349.
XX
PR   07-DEC-2000; 2000US-0254648P.
PR   13-MAR-2001; 2001US-0275688P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (HYSE-) HYSEQ INC.
XX
PI   Escobedo J,  Garcia PD,  Kassam A,  Lamson G,  Drmanac R;
PI   Crkvenjakov R,  Dickson M,  Drmanac S,  Labat I,  Leshkowitz D,  Kita D;
PI   Garcia V,  Jones WL,  Stache-Crain B,  Scott EM;
XX
DR   WPI; 2002-557824/59.
XX
PT   New genes and gene products isolated from human prostate, useful for
PT   treating or diagnosing tumor or cancer (e.g. prostate cancer or breast
PT   cancer), or as vaccines for treating or preventing these diseases.
XX
PS   Claim 1; SEQ ID NO 494; 186pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated polynucleotide comprising any of
CC   1477 sequences or its fragment, degenerate variant, antisense or
CC   complement. The polynucleotides and gene products are useful for treating
CC   or diagnosing tumour or cancer (e.g. prostate cancer, breast cancer, lung
CC   cancer or medullary carcinoma) in a subject (e.g. cattle, dogs, cats,
CC   rabbits, horse or human). The polynucleotides and polypeptides are also
CC   useful as vaccines for treating or preventing these diseases. The
CC   polynucleotides are useful for gene therapy. The present sequence is that
CC   of one of a group of polynucleotides (ABQ88745-ABQ90015) disclosed
CC   electronically as sequences of the invention. However only 1271
CC   polynucleotide sequences are given, whereas 1477 polynucleotides and 91
CC   proteins are claimed. Note: The sequence data for this patent did not
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequence
XX
SQ   Sequence 835 BP; 209 A; 183 C; 185 G; 232 T; 0 U; 26 Other;

  Query Match           25.6%;  Score 27.4;  DB 6;  Length 835;
  Best Local Similarity 24.3%;  Pred. No. 92;
  Matches   25;  Conservative    9;  Mismatches   69;  Indels    0;  Gaps    0;

Qy          2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCU 61
              |||||      : | |         :    :       ||||  :| |||| :     |:
Db        311 GAAAACTCTTGTTAGACATGATGTTTCCTATACAGAAGAAAAGATACAAACTTGATTGCT 370
```

```
Qy          62 UNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAG 104
               : ||           ||| |:             : ||
Db         371 TCAGTAAAAACACTTTTAGAGATTGTGTTTCCTTTTTTGTGAG 413


RESULT 13
ADB83094
ID   ADB83094 standard; cDNA; 929 BP.
XX
AC   ADB83094;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Human cDNA sequence useful for the treatment of cancer (SeqID 1307).
XX
KW   human; prostate; cancer; cytostatic; gene therapy; vaccine;
KW   immune response; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2003050236-A2.
XX
PD   19-JUN-2003.
XX
PF   04-SEP-2002; 2002WO-US028214.
XX
PR   07-DEC-2001; 2001US-00012697.
XX
PA   (CHIR ) CHIRON CORP.
PA   (HYSE-) HYSEQ INC.
XX
PI   Escobedo J,  Garcia PD,  Kassam A,  Lamson G,  Drmanac R;
PI   Crkvenjakov R,  Dickson M,  Drmanac S,  Labat I,  Leshkowitz D,  Kita D;
PI   Garcia V,  Jones LW,  Stache-Crain B,  Scott EM;
XX
DR   WPI; 2003-513972/48.
XX
PT   New polynucleotides derived from human prostate, useful for modulating
PT   immune response to prevent or treat cancer.
XX
PS   Claim 1; SEQ ID NO 1307; 188pp; English.
XX
CC   This invention relates to novel isolated polynucleotides of human origin,
CC   particularly isolated from the human prostate. Specifically, it refers to
CC   the diagnostics and therapeutics comprising these novel human
CC   polynucleotides, and includes the derived probes, antisense
CC   oligonucleotides and antibodies thereof. The identification of these
CC   human prostate genes that can inhibit tumour growth is useful for
CC   understanding the progression and nature of complex diseases such as
CC   cancer, and hence they are important in the drug discovery process. The
CC   present invention describes these polynucleotides and encoded
CC   polypeptides as exhibiting cytostatic activity, and through gene therapy
CC   and/ or vaccines they can be used to modulate the immune response for the
CC   prevention or treatment of cancers, particularly of the prostate, but
CC   also for breast, lung and colon cancer. This polynucleotide sequence is a
CC   human cDNA sequence useful for the treatment of cancer, used in an
CC   exemplification of the invention. NOTE: These sequences are not given in
CC   the specification but are provided on the WIPO website.
XX
SQ   Sequence 929 BP; 243 A; 207 C; 211 G; 244 T; 0 U; 24 Other;

  Query Match             25.6%;  Score 27.4;  DB 8;  Length 929;
  Best Local Similarity   24.3%;  Pred. No. 93;
  Matches   25;  Conservative    9;  Mismatches   69;  Indels    0;  Gaps    0;

Qy           2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNNAAAAANUANAAACAUNNNNNCU 61
               |||||      : | |         :   :       ||||  :| |||| :     |:
Db         237 GAAAACTCTTGTTAGACATGATGTTTCCTATACAGAAGAAAAGATACAAACTTGATTGCT 296

Qy          62 UNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAG 104
               : ||           ||| |:             : ||
Db         297 TCAGTAAAAACACTTTTAGAGATTGTGTTTCCTTTTTTGTGAG 339


RESULT 14
AAS43104
ID   AAS43104 standard; DNA; 325791 BP.
XX
AC   AAS43104;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Human Oestrogen receptor beta gene.
XX
KW   Human; Oestrogen receptor beta; ERbeta; ds; SNP; chromosome 6q.25.1;
KW   single nucleotide polymorphism; cardiovascular disease;
KW   autoimmune disease; systemic lupus erythematosus; arthritis; rheumatism;
KW   osteoarthritis; osteoporosis; breast cancer; endometrial cancer.
XX
OS   Homo sapiens.
XX
PN   WO200162793-A2.
XX
PD   30-AUG-2001.
XX
PF   20-FEB-2001; 2001WO-US005360.
XX
PR   22-FEB-2000; 2000US-0183755P.
PR   24-JAN-2001; 2001US-00768185.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Kalush F,  Cassel MJ,  Hwang SS,  Winn-Deen ES;
XX
DR   WPI; 2001-582041/65.
DR   P-PSDB; AAU27322.
XX
PT   Estrogen receptor gene and protein polymorphisms useful for diagnosis of
PT   individuals at risk of developing bone disorders.
XX
PS   Example 2; Fig 1; 245pp; English.
XX
CC   The invention relates to a novel isolated peptide comprising or
CC   consisting of an amino acid sequence selected from an amino acid sequence
CC   of a variant oestrogen receptor protein (e.g. ERbeta), or a fragment of
CC   10 amino acids), antibodies against them, nucleic acids encoding them
CC   (including vectors for transforming cells). The gene for human ERbeta is
CC   located on chromosome 6q.25.1. The variants are encoded by single
CC   nucleotide polymorphisms (SNP). The variant peptides and proteins can be
CC   used in assays to determine the biological activity of the protein, to
CC   raise antibodies, as a reagent in assays designed to quantitatively
CC   determine levels of the protein in biological fluids, to identify
CC   compounds that modulate receptor activity and to screen compounds for the
CC   ability to stimulate or inhibit interaction between the receptor protein
CC   and a target molecule that normally interacts with the receptor protein
CC   e.g. oestrogen. The antibody can be used to isolate the protein, to
CC   assess expression in disease states e.g. cardiovascular disease and
CC   autoimmune disease (e.g. systemic lupus erythematosus, arthritis,
CC   rheumatism and osteoarthritis), osteoporosis, breast cancer and
CC   endometrial cancer. In addition the antibodies can be used in
CC   pharmacogenomic analysis and inhibiting protein function, e.g. blocking
CC   the binding of the oestrogen receptor protein to a binding partner such
CC   as a ligand. The nucleic acids encoding the proteins can be used as
CC   probes, primers, chemical intermediates and in biological assays. The
CC   present sequence is the human ERbeta gene
XX
SQ   Sequence 325791 BP; 94098 A; 68292 C; 67970 G; 95431 T; 0 U; 0 Other;

  Query Match             25.2%;  Score 27;  DB 4;  Length 325791;
  Best Local Similarity   20.8%;  Pred. No. 2.5e+02;
  Matches   21;  Conservative   12;  Mismatches   68;  Indels    0;  Gaps    0;

Qy           3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNNAAAAANUANAAACAUNNNNNCUU 62
               ||  :       |:|:|       |:   :::       |||  |  | |||| :      ::
Db      205486 AATTTGACCTGCATATGCATACTCTATCTTTGGAAGAAAATACAACAAACTTGTAACATT 205545
```

```
Qy          63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUA 103
               |                 || ||                  ::|
Db      205546 GATTGCTTCTGTGGAGAGGAACTGATAAGAGAAAAATTTTA 205586
```

RESULT 15
ABL32266/c
ID ABL32266 standard; DNA; 6419 BP.
XX
AC ABL32266;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 239.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW ds.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP007537.
XX
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.
XX
PS Claim 1; SEQ ID NO 239; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention
XX
SQ Sequence 6419 BP; 1723 A; 53 C; 1678 G; 2965 T; 0 U; 0 Other;

```
  Query Match             25.0%;  Score 26.8;  DB 6;  Length 6419;
  Best Local Similarity   24.0%;  Pred. No. 1.7e+02;
  Matches   25;  Conservative    9;  Mismatches   70;  Indels    0;  Gaps    0;

Qy           3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
               ||||      : |:|       |     ::      |||||    ||| |:     |::
Db        1272 AAAAAATAAATTATAATTAATTTCCCATCTTACAAATAAAAAAACTAAAAATCGAAACTT 1213

Qy          63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAGCA 106
               |              | |||:                 :| ||
Db        1212 AAAAAACAACCTACCCAAAAATCACCCAACAAACTATTCTAACA 1169
```

Search completed: March 8, 2004, 04:45:33
Job time : 310 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 8, 2004, 04:31:47 ; Search time 60 Seconds
(without alignments)
989.662 Million cell updates/sec

Title: US-09-646-651C-2
Perfect score: 107
Sequence: 1 ggaaaaunnnnnunauaugn.....nnnnnnnnnnnnnnuuagcag 107

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	29.4	27.5	1664976	4	US-08-916-421B-1	Sequence 1, Appli
	2	26	24.3	318	4	US-09-328-352-642	Sequence 642, App
c	3	25.8	24.1	934	3	US-08-842-306B-5	Sequence 5, Appli
c	4	25.8	24.1	934	3	US-08-838-973B-5	Sequence 5, Appli
c	5	25.8	24.1	934	4	US-08-771-212A-5	Sequence 5, Appli
c	6	24.6	23.0	499	4	US-09-621-976-13988	Sequence 13988, A
	7	24.6	23.0	3552	4	US-09-643-597-126	Sequence 126, App
	8	24.6	23.0	3552	4	US-09-480-884A-126	Sequence 126, App
	9	24.6	23.0	3552	4	US-09-542-615A-126	Sequence 126, App
	10	24.6	23.0	3552	4	US-09-606-421B-126	Sequence 126, App
	11	24.6	23.0	3552	4	US-09-221-107-126	Sequence 126, App
c	12	24.6	23.0	5652	4	US-09-601-198-75	Sequence 75, Appl
c	13	24.4	22.8	398	4	US-09-401-064-281	Sequence 281, App
c	14	24.4	22.8	5614	4	US-08-956-171E-99	Sequence 99, Appl
	15	24.4	22.8	5915	4	US-10-204-708-77	Sequence 77, Appl
c	16	24.2	22.6	8093	4	US-10-204-708-32	Sequence 32, Appl
c	17	24.2	22.6	640681	4	US-09-790-988-1	Sequence 1, Appli
	18	24	22.4	927	4	US-09-134-000C-3089	Sequence 3089, Ap
c	19	24	22.4	6583	4	US-10-204-708-25	Sequence 25, Appl
	20	24	22.4	11131	4	US-10-204-708-28	Sequence 28, Appl
	21	23.8	22.2	65792	4	US-09-596-002-31	Sequence 31, Appl
	22	23.6	22.1	1474	4	US-08-858-207A-144	Sequence 144, App
c	23	23.6	22.1	2418	4	US-09-601-198-61	Sequence 61, Appl
	24	23.6	22.1	8703	4	US-08-961-527-177	Sequence 177, App
	25	23.4	21.9	2319	2	US-08-966-389-5	Sequence 5, Appli
	26	23.4	21.9	2319	2	US-09-103-509-5	Sequence 5, Appli
	27	23.4	21.9	2319	2	US-09-102-644-5	Sequence 5, Appli
	28	23.4	21.9	2319	2	US-09-218-032-5	Sequence 5, Appli
c	29	23.4	21.9	2675	4	US-09-636-791A-12	Sequence 12, Appl
	30	23.4	21.9	3345	2	US-08-966-389-8	Sequence 8, Appli
	31	23.4	21.9	3345	2	US-09-103-509-8	Sequence 8, Appli
	32	23.4	21.9	3345	2	US-09-102-644-8	Sequence 8, Appli
	33	23.4	21.9	3345	2	US-09-218-032-8	Sequence 8, Appli
c	34	23.4	21.9	9347	4	US-10-204-708-35	Sequence 35, Appl
	35	23.4	21.9	1664976	4	US-08-916-421B-1	Sequence 1, Appli
	36	23.2	21.7	711	1	US-08-307-499-49	Sequence 49, Appl
	37	23.2	21.7	711	3	US-09-299-268-49	Sequence 49, Appl
c	38	23.2	21.7	727	1	US-07-885-970A-6	Sequence 6, Appli
c	39	23.2	21.7	727	1	US-08-298-687A-6	Sequence 6, Appli
c	40	23.2	21.7	727	1	US-08-530-797-5	Sequence 5, Appli
c	41	23.2	21.7	727	1	US-08-298-829-6	Sequence 6, Appli
c	42	23.2	21.7	727	2	US-08-787-335-5	Sequence 5, Appli
c	43	23.2	21.7	1128	4	US-09-543-681A-3553	Sequence 3553, Ap
c	44	23.2	21.7	10144	4	US-10-204-708-94	Sequence 94, Appl
c	45	23.2	21.7	14176	1	US-08-307-499-1	Sequence 1, Appli

ALIGNMENTS

```
RESULT 1
US-08-916-421B-1/c
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
;  APPLICANT: Bult et al.
;  TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococ
; Patent No. 6503729
;  TITLE OF INVENTION:  jannaschii
;  FILE REFERENCE: PB275
;  CURRENT APPLICATION NUMBER: US/08/916,421B
;  CURRENT FILING DATE:  1997-08-22
;  PRIOR APPLICATION NUMBER: US 60/024,428
;  PRIOR FILING DATE: 1996-08-22
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1664976
;   TYPE: DNA
;   ORGANISM: Methanococcus jannaschii
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (28222)..(28222)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (28257)..(28258)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (84773)..(84773)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (84808)..(84808)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (84812)..(84812)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98120)..(98120)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98159)..(98159)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98239)..(98239)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98266)..(98266)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98343)..(98343)
```

; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g

; NAME/KEY: misc_feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1349491)..(1349491)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1470091)..(1470091)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1569020)..(1569020)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1602912)..(1602912)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1603734)..(1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1637998)..(1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1664854)..(1664855)
; OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

Query Match 27.5%; Score 29.4; DB 4; Length 1664976;
Best Local Similarity 24.8%; Pred. No. 4;
Matches 26; Conservative 10; Mismatches 69; Indels 0; Gaps 0;

```
Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              ||||:     : | | |       |:   :::      ||||| :| |||||:      :
Db     573299 AAAATTAGTATTAGAAGCTATAACTTATTTTAATGCCAAAAAATATAAACATTTCTGGTG 573240

Qy         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUAGCAG 107
                |              ||  |:             | ||
Db     573239 AAATAATGACAGGCAAAGTTATCTTAGTAGGAGCAGGACCGGGAG 573195
```

RESULT 2
US-09-328-352-642
; Sequence 642, Application US/09328352
; Patent No. 6562958

GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 642
LENGTH: 318
TYPE: DNA
ORGANISM: Acinetobacter baumannii
US-09-328-352-642

Query Match 24.3%; Score 26; DB 4; Length 318;
Best Local Similarity 26.2%; Pred. No. 9.5;
Matches 22; Conservative 10; Mismatches 52; Indels 0; Gaps 0;

Qy 1 GGAAAAUNNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNC 60
Db 34 GCAAAATATATACAAATTGAAATATCTTTTTTAGCACTTAAATATTTAAAAAGTGGTAAC 93

Qy 61 UUNAGNNNNNNNNNNNNAGAAAU 84
Db 94 TTAATTAATAAAACAAAAAGAAAT 117

RESULT 3
US-08-842-306B-5/c
Sequence 5, Application US/08842306B
Patent No. 6271197
GENERAL INFORMATION:
APPLICANT: Berlin, Vivian
Levin, David
Ohya, Yoshikazu
Damagnez, Veronique
Smith, Susan
TITLE OF INVENTION: ASSAYS AND REAGENTS FOR IDENTIFYING
ANTI-FUNGAL AGENTS, AND USES RELATED THERETO
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPad
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/842,306B
FILING DATE: 23-Apr-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/771,212
FILING DATE: 20-DEC-1996
APPLICATION NUMBER: US 08/631,319
FILING DATE: 11-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MIV-074.04
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 934 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 260..832
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-08-842-306B-5

Query Match 24.1%; Score 25.8; DB 3; Length 934;
Best Local Similarity 22.3%; Pred. No. 13;
Matches 23; Conservative 10; Mismatches 70; Indels 0; Gaps 0;

Qy 4 AAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUUN 63
Db 243 AATTATTTATGAGCTGAATACTTTTCTCTTGTGCTAAAAAAAGAAAAAACTGAAAGCAAA 184

Qy 64 AGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAGCA 106
Db 183 AGTGAAAACTAAAGTAAAAATAACAATAAAAATAGATTTAGGA 141

RESULT 4
US-08-838-973B-5/c
Sequence 5, Application US/08838973B
Patent No. 6277564
GENERAL INFORMATION:
APPLICANT: Berlin, Vivian
Damagnez, Veronique
Smith, Susan
TITLE OF INVENTION: ASSAYS AND REAGENTS FOR IDENTIFYING
ANTI-FUNGAL AGENTS, AND USES RELATED THERETO
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/838,973B
FILING DATE: 23-Apr-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/631,319
FILING DATE: 10-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MIV-074.05
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 934 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 260..832
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-08-838-973B-5

Query Match 24.1%; Score 25.8; DB 3; Length 934;
Best Local Similarity 22.3%; Pred. No. 13;
Matches 23; Conservative 10; Mismatches 70; Indels 0; Gaps 0;

```
Qy          4 AAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUUN 63
              || :     : |  :|       :    ::      |||||  | |||  :     |
Db        243 AATTATTTATGAGCTGAATACTTTTCTCTTGTGCTAAAAAAAGAAAAAACTGAAAGCAAA 184

Qy         64 AGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAGCA 106
              ||             | |||:             ::|| |
Db        183 AGTGAAAACTAAAGTAAAAATAACAATAAAAATAGATTTAGGA 141


RESULT 5
US-08-771-212A-5/c
; Sequence 5, Application US/08771212A
; Patent No. 6455281
; GENERAL INFORMATION:
;  APPLICANT: Berlin, V.
;  APPLICANT:  Damagnez, V.
;  APPLICANT:  Smith, S.
;  TITLE OF INVENTION: NUCLEIC ACIDS FOR IDENTIFYING ANTI-FUNGAL AGENTS, AND USES RELATE
;  TITLE OF INVENTION:  THERETO
;  FILE REFERENCE: GPCI-P03-074
;  CURRENT APPLICATION NUMBER: US/08/771,212A
;  CURRENT FILING DATE:  1996-12-20
;  PRIOR APPLICATION NUMBER: 08/631,319
;  PRIOR FILING DATE: 1996-04-10
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
;   LENGTH: 934
;   TYPE: DNA
;   ORGANISM: Candida sp
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (260)..(832)
;   OTHER INFORMATION:
US-08-771-212A-5

  Query Match             24.1%;  Score 25.8;  DB 4;  Length 934;
  Best Local Similarity   22.3%;  Pred. No. 13;
  Matches   23;  Conservative   10;  Mismatches   70;  Indels    0;  Gaps    0;

Qy          4 AAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUUN 63
              || :     : |  :|       :    ::      |||||  | |||  :     |
Db        243 AATTATTTATGAGCTGAATACTTTTCTCTTGTGCTAAAAAAAGAAAAAACTGAAAGCAAA 184

Qy         64 AGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAGCA 106
              ||             | |||:             ::|| |
Db        183 AGTGAAAACTAAAGTAAAAATAACAATAAAAATAGATTTAGGA 141


RESULT 6
US-09-621-976-13988/c
; Sequence 13988, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 13988
;   LENGTH: 499
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-621-976-13988

  Query Match             23.0%;  Score 24.6;  DB 4;  Length 499;
  Best Local Similarity   21.9%;  Pred. No. 27;
  Matches   23;  Conservative   10;  Mismatches   72;  Indels    0;  Gaps    0;

Qy          1 GGAAAAUNNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNC 60
              | | ||      : |:|:|       :   :          || | :  |||  |:     |
Db        423 GCATAACCTGCATCATATGGCTGTCTTGCTTAGCTGCAAGAATAATGAAAATATTCTACC 364

Qy         61 UUNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAGC 105
              :  |              |||||                ::| |
Db        363 TCAATTTTATGGAGCATAAGAAAATTTTTTGAGGAAGGACTTAAC 319


RESULT 7
US-09-643-597-126
; Sequence 126, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  McNeill, Patricia D.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.455C11
;  CURRENT APPLICATION NUMBER: US/09/643,597
;  CURRENT FILING DATE:  2000-08-21
;  NUMBER OF SEQ ID NOS:  369
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 126
;   LENGTH: 3552
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-643-597-126

  Query Match             23.0%;  Score 24.6;  DB 4;  Length 3552;
  Best Local Similarity   27.2%;  Pred. No. 38;
  Matches   22;  Conservative    8;  Mismatches   51;  Indels    0;  Gaps    0;

Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              || |:     :  :|:|      |:     :          |||  | |||||       ::
Db       3081 AATATCACATTATTATGTATTCACTTTAAGTGATAGTTTAAAAAATAAACAAGAAATATT 3140

Qy         63 NAGNNNNNNNNNNNNNAGAAA 83
              ||             |||||
Db       3141 GAGTATCACTATGTGAAGAAA 3161


RESULT 8
US-09-480-884A-126
; Sequence 126, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Hosken, Nancy A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C6
;  CURRENT APPLICATION NUMBER: US/09/480,884A
;  CURRENT FILING DATE:  2001-08-27
;  NUMBER OF SEQ ID NOS: 330
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 126
;   LENGTH: 3552
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-480-884A-126
```

```
 Query Match             23.0%;  Score 24.6;  DB 4;  Length 3552;
 Best Local Similarity   27.2%;  Pred. No. 38;
 Matches   22;  Conservative    8;  Mismatches   51;  Indels    0;  Gaps    0;

y          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
             || |:     :  :|:|      |:     :        |||  | |||||       ::
b       3081 AATATCACATTATTATGTATTCACTTTAAGTGATAGTTTAAAAAATAAACAAGAAATATT 3140

y         63 NAGNNNNNNNNNNNNNAGAAA 83
             ||             |||||
b       3141 GAGTATCACTATGTGAAGAAA 3161


ESULT 9
S-09-542-615A-126
 Sequence 126, Application US/09542615A
 Patent No. 6518256
 GENERAL INFORMATION:
  APPLICANT:  Wang, Tongtong
  APPLICANT:  Fan, Liqun
  APPLICANT:  Kalos, Michael D.
  APPLICANT:  Bangur, Chaitanya S.
  APPLICANT:  Hosken, Nancy A.
  APPLICANT:  Fanger, Gary R.
  TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THERAPY
  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
  FILE REFERENCE:  210121.455C8
  CURRENT APPLICATION NUMBER: US/09/542,615A
  CURRENT FILING DATE:  2000-04-14
  NUMBER OF SEQ ID NOS:  350
  SOFTWARE:  FastSEQ for Windows Version 3.0
 SEQ ID NO 126
    LENGTH: 3552
    TYPE: DNA
    ORGANISM: Homo sapien
S-09-542-615A-126

 Query Match             23.0%;  Score 24.6;  DB 4;  Length 3552;
 Best Local Similarity   27.2%;  Pred. No. 38;
 Matches   22;  Conservative    8;  Mismatches   51;  Indels    0;  Gaps    0;

y          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
             || |:     :  :|:|      |:     :        |||  | |||||       ::
b       3081 AATATCACATTATTATGTATTCACTTTAAGTGATAGTTTAAAAAATAAACAAGAAATATT 3140

y         63 NAGNNNNNNNNNNNNNAGAAA 83
             ||             |||||
b       3141 GAGTATCACTATGTGAAGAAA 3161


ESULT 10
S-09-606-421B-126
 Sequence 126, Application US/09606421B
 Patent No. 6531315
 GENERAL INFORMATION:
  APPLICANT: Wang, Tongtong
  APPLICANT:  Fan, Liqun
  APPLICANT:  Kalos, Michael D.
  APPLICANT:  Bangur, Chaitanya S.
  APPLICANT:  Hosken, Nancy
  APPLICANT:  Fanger, Gary R.
  APPLICANT:  Li, Samuel X.
  APPLICANT:  Wang, Aijun
  APPLICANT:  Skeiky, Yasir A.W.
  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
  FILE REFERENCE: 210121.455C9
  CURRENT APPLICATION NUMBER: US/09/606,421B
  CURRENT FILING DATE:  2000-06-28
  NUMBER OF SEQ ID NOS: 358
  SOFTWARE: FastSEQ for Windows Version 3.0
 SEQ ID NO 126
;   LENGTH: 3552
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-606-421B-126

  Query Match             23.0%;  Score 24.6;  DB 4;  Length 3552;
  Best Local Similarity   27.2%;  Pred. No. 38;
  Matches   22;  Conservative    8;  Mismatches   51;  Indels    0;  Gaps    0;

Qy         3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
             || |:     :  :|:|      |:     :        |||  | |||||       ::
Db      3081 AATATCACATTATTATGTATTCACTTTAAGTGATAGTTTAAAAAATAAACAAGAAATATT 3140

Qy        63 NAGNNNNNNNNNNNNNAGAAA 83
             ||             |||||
Db      3141 GAGTATCACTATGTGAAGAAA 3161


RESULT 11
US-09-221-107-126
; Sequence 126, Application US/09221107
; Patent No. 6660838
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
;  FILE REFERENCE: 210121.455C2
;  CURRENT APPLICATION NUMBER: US/09/221,107
;  CURRENT FILING DATE: 1998-12-22
;  NUMBER OF SEQ ID NOS: 161
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 126
;   LENGTH: 3552
;   TYPE: DNA
;   ORGANISM: Human
US-09-221-107-126

  Query Match             23.0%;  Score 24.6;  DB 4;  Length 3552;
  Best Local Similarity   27.2%;  Pred. No. 38;
  Matches   22;  Conservative    8;  Mismatches   51;  Indels    0;  Gaps    0;

Qy         3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
             || |:     :  :|:|      |:     :        |||  | |||||       ::
Db      3081 AATATCACATTATTATGTATTCACTTTAAGTGATAGTTTAAAAAATAAACAAGAAATATT 3140

Qy        63 NAGNNNNNNNNNNNNNAGAAA 83
             ||             |||||
Db      3141 GAGTATCACTATGTGAAGAAA 3161


RESULT 12
US-09-601-198-75/c
; Sequence 75, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
;  APPLICANT: Cassell, Gail H.
;  APPLICANT:  Chen, Ellson Y.
;  APPLICANT:  Glass, Jennifer S.
;  APPLICANT:  Glass, John I.
;  APPLICANT:  Heiner, Cheryl R.
;  APPLICANT:  Lefkowitz, Elliot
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
;  TITLE OF INVENTION:  UREALYTICUM
;  FILE REFERENCE: UAB-13452/22
;  CURRENT APPLICATION NUMBER: US/09/601,198
;  CURRENT FILING DATE:  2000-12-08
;  PRIOR APPLICATION NUMBER: 60/073,189
;  PRIOR FILING DATE: 1998-01-30
;  NUMBER OF SEQ ID NOS: 181
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 75
;   LENGTH: 5652
;   TYPE: DNA
;   ORGANISM: Ureaplasma urealyticum
```

```
US-09-601-198-75

  Query Match             23.0%;  Score 24.6;  DB 4;  Length 5652;
  Best Local Similarity   21.0%;  Pred. No. 42;
  Matches   22;  Conservative   11;  Mismatches   72;  Indels    0;  Gaps    0;

Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              ||||:      | |:       :   : :      | | | :| ||| |       ::
Db        413 AAAATAATGGCAAAATAGAATATATCGATGTTCACACAGATAATACAAATAATTTATTTT 354

Qy         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAGCAG 107
                          |||:                  ::|| ||
Db        353 ATCTAAAACGTCAATCAGACATTAATTCTAAATGAACATTAGAAG 309


RESULT 13
US-09-401-064-281/c
; Sequence 281, Application US/09401064
; Patent No. 6623923
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT: Lodes, Michael J.
;  APPLICANT: Secrist, Heather
;  APPLICANT: Benson, Darin R.
;  APPLICANT: Meagher, Madeline Joy
;  APPLICANT: Stolk, John A.
;  APPLICANT: Wang, Tongtong
;  TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
;  TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
;  FILE REFERENCE: 210121.471C2
;  CURRENT APPLICATION NUMBER: US/09/401,064
;  CURRENT FILING DATE: 1999-09-22
;  NUMBER OF SEQ ID NOS: 371
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 281
;   LENGTH: 398
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-401-064-281

  Query Match             22.8%;  Score 24.4;  DB 4;  Length 398;
  Best Local Similarity   23.2%;  Pred. No. 30;
  Matches   22;  Conservative    9;  Mismatches   64;  Indels    0;  Gaps    0;

Qy         13 UNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUUNAGNNNNNNN 72
              :  : :      |     ::      | ||  | |||||:       : ||
Db        204 TGTTTTCCTATTACCATTATTAAAAGTCAGAAGGAGAAACATGAAGAACTAAGTGACAAG 145

Qy         73 NNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAGCAG 107
                    |||||:              : |||||
Db        144 ATAATGAGAAATATTTGAGGCTGTAAAATAAGCAG 110


RESULT 14
US-08-956-171E-99/c
, Sequence 99, Application US/08956171E
; Patent No. 6593114
;   GENERAL INFORMATION:
;        APPLICANT: Charles Kunsch
;                   Gil H. Choi
;                   Patrick S. Dillon
;                   Craig A. Rosen
;                   Steven C. Barash
;                   Michael R. Fannon
;        TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;        NUMBER OF SEQUENCES: 5256
;        CORRESPONDENCE ADDRESS:
,             ADDRESSEE: Human Genome Sciences, Inc.
,             STREET: 9410 Key West Avenue
,             CITY: Rockville
,             STATE: Maryland
,             COUNTRY: USA
,             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MSDOS version 6.2
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/956,171E
;             FILING DATE: 20-Oct-1997
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/009,861
;             FILING DATE: January 5, 1996
;             APPLICATION NUMBER: 08/781,986
;             FILING DATE: January 3, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Mark J. Hyman
;             REGISTRATION NUMBER: 46,789
;             REFERENCE/DOCKET NUMBER: PB248P1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (240) 314-1224
;             TELEFAX: (301) 309-8439
;   INFORMATION FOR SEQ ID NO: 99:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 5614 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 99:
US-08-956-171E-99

  Query Match             22.8%;  Score 24.4;  DB 4;  Length 5614;
  Best Local Similarity   22.0%;  Pred. No. 48;
  Matches   22;  Conservative    9;  Mismatches   69;  Indels    0;  Gaps    0;

Qy          4 AAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUUN 63
              |||:      : | | |       :    :      ||   | :| ||| |       |::
Db       1552 AAATGGATTTTAAAAGATAGCAATTAAAGTAGCGTAAACCACTAAAAAAAACTTTACTTA 1493

Qy         64 AGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUA 103
              ||             ||||:                 :|
Db       1492 AGCCTTAATATTGCGTGAAATATAATAAATAAATGATATA 1453


RESULT 15
US-10-204-708-77
; Sequence 77, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
;  APPLICANT: OLEK, Alexander
;  APPLICANT:  PIEPENBROCK, Christian
;  APPLICANT:  BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
;  TITLE OF INVENTION:  by Assessing DNA Methylation
;  FILE REFERENCE: 5013.1012
;  CURRENT APPLICATION NUMBER: US/10/204,708
;  CURRENT FILING DATE:  2003-05-06
;  PRIOR APPLICATION NUMBER: PCT/EP01/03971
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: DE 10019058.8
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: DE 10019173.8
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: DE 10032529.7
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: DE 10043826.1
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 77
;   LENGTH: 5915
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
```

US-10-204-708-77

Query Match 22.8%; Score 24.4; DB 4; Length 5915;
Best Local Similarity 19.0%; Pred. No. 48;
Matches 19; Conservative 12; Mismatches 69; Indels 0; Gaps 0;

```
Qy          5 AAUNNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNNAAAAANUANAAACAUNNNNNCUUNA 64
              ||:     : |:|:|       :   :::       ||||| :| ||| |:
Db       3458 AATTTGGTTAATATGGTAAAATTTTGTTTTTATTAAAAAAATATAAAAATTAGTTAGGCG 3517

Qy         65 GNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUAG 104
                        | ||:                   : |
Db       3518 TGGTGGTAGGCGTTTGTAATTTTAGTTATACGGGAGGTTG 3557
```

Search completed: March 8, 2004, 06:01:49
Job time : 65 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

)M nucleic - nucleic search, using sw model

:un on: March 8, 2004, 04:40:28 ; Search time 1220 Seconds
(without alignments)
320.775 Million cell updates/sec

:itle: US-09-646-651C-2
)erfect score: 107
)equence: 1 ggaaaaunnnnnunauaugn.....nnnnnnnnnnnnnuuagcag 107

)coring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

)earched: 2421054 seqs, 1828716029 residues

)otal number of hits satisfying chosen parameters: 4842108

)inimum DB seq length: 0
)aximum DB seq length: 2000000000

)ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

)atabase : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

esult No.	Score	% Query Match	Length	DB	ID	Description
1	27.4	25.6	524	15	US-10-012-697-280	Sequence 280, App
2	27.4	25.6	761	15	US-10-012-697-493	Sequence 493, App
3	27.4	25.6	769	15	US-10-012-697-418	GENERAL INFORMATI
4	27.4	25.6	789	15	US-10-012-697-415	Sequence 415, App
5	27.4	25.6	929	15	US-10-012-697-1307	Sequence 1307, Ap
c 6	26.8	25.0	6419	14	US-10-311-455-239	Sequence 239, App
c 7	26.8	25.0	13814	14	US-10-311-455-1166	Sequence 1166, Ap
c 8	26.8	25.0	3673778	14	US-10-312-841-2	Sequence 2, Appli
c 9	26.6	24.9	487	15	US-10-027-632-323248	Sequence 323248,
c 10	26.6	24.9	488	15	US-10-027-632-324208	Sequence 324208,
11	26.6	24.9	634	15	US-10-027-632-8286	Sequence 8286, Ap
12	26.6	24.9	10034	14	US-10-311-455-1858	Sequence 1858, Ap
c 13	26.2	24.5	390	10	US-09-814-353-17808	Sequence 17808, A
c 14	26.2	24.5	3673778	14	US-10-312-841-1	Sequence 1, Appli
15	26	24.3	511	12	US-10-424-599-85840	Sequence 85840, A
c 16	26	24.3	5504	12	US-10-221-613-270	Sequence 270, App
17	26	24.3	7977	14	US-10-311-455-2062	Sequence 2062, Ap
18	26	24.3	3673778	14	US-10-312-841-1	Sequence 1, Appli
19	25.8	24.1	553	9	US-09-560-863-928	Sequence 928, App
20	25.8	24.1	585	15	US-10-027-632-251034	Sequence 251034,
c 21	25.8	24.1	934	9	US-09-945-249-5	Sequence 5, Appli
c 22	25.8	24.1	1031	14	US-10-093-524-9	Sequence 9, Appli
c 23	25.8	24.1	1386	12	US-10-282-122A-15352	Sequence 15352, A
24	25.8	24.1	1566	14	US-10-081-051-59	Sequence 59, Appl
25	25.8	24.1	1905	12	US-10-282-122A-17125	Sequence 17125, A
26	25.8	24.1	4544	14	US-10-081-051-58	Sequence 58, Appl
c 27	25.8	24.1	7461	9	US-09-764-860-1135	Sequence 1135, Ap
c 28	25.8	24.1	7461	9	US-09-764-904-91	Sequence 91, Appl
c 29	25.8	24.1	7461	14	US-10-091-548-91	Sequence 91, Appl
c 30	25.8	24.1	7461	14	US-10-074-095-1135	Sequence 1135, Ap
c 31	25.8	24.1	7461	15	US-10-212-872-1135	Sequence 1135, Ap
c 32	25.8	24.1	32174	9	US-09-908-711-158	Sequence 158, App
c 33	25.8	24.1	32174	9	US-09-764-860-1134	Sequence 1134, Ap
c 34	25.8	24.1	32174	9	US-09-764-877-2645	Sequence 2645, Ap
c 35	25.8	24.1	32174	9	US-09-860-670-232	Sequence 232, App
c 36	25.8	24.1	32174	9	US-09-764-904-90	Sequence 90, Appl
c 37	25.8	24.1	32174	10	US-09-764-891-6480	Sequence 6480, Ap
38	25.8	24.1	32174	10	US-09-764-891-10135	Sequence 10135, A
c 39	25.8	24.1	32174	10	US-09-764-891-10179	Sequence 10179, A
c 40	25.8	24.1	32174	14	US-10-091-548-90	Sequence 90, Appl
c 41	25.8	24.1	32174	14	US-10-074-095-1134	Sequence 1134, Ap
c 42	25.8	24.1	32174	15	US-10-212-872-1134	Sequence 1134, Ap
43	25.8	24.1	32174	15	US-10-074-024-609	Sequence 609, App
c 44	25.8	24.1	32174	15	US-10-227-646-232	Sequence 232, App
c 45	25.8	24.1	32174	15	US-10-242-515-2645	Sequence 2645, Ap

ALIGNMENTS

```
RESULT 1
US-10-012-697-280
; Sequence 280, Application US/10012697
; Publication No. US20030215803A1
; GENERAL INFORMATION:
;  APPLICANT: Escobedo, Jaime
;  APPLICANT:  Garcia, Pablo Dominguez
;  APPLICANT:  Kassam, Altaf
;  APPLICANT:  Lamson, George
;  APPLICANT:  Scott, Beth
;  APPLICANT:  Drmanac, Radoje
;  APPLICANT:  Crkvenjakov, Radomir
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Drmanac, Snezana
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Leshkowitz, Dena
;  APPLICANT:  Kita, David
;  APPLICANT:  Garcia, Veronica
;  APPLICANT:  Jones, Lee William
;  APPLICANT:  Stache-Crain, Birgit
;  TITLE OF INVENTION: HUMAN GENES AND GENE EXPRESSION PRODUCTS
;  TITLE OF INVENTION:  ISOLATED FROM HUMAN PROSTATE
;  FILE REFERENCE: 2300-16252
;  CURRENT APPLICATION NUMBER: US/10/012,697
;  CURRENT FILING DATE:  2003-01-21
;  PRIOR APPLICATION NUMBER: 60/254,648
;  PRIOR FILING DATE: 2000-12-07
;  PRIOR APPLICATION NUMBER: 60/275,668
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 1568
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 280
;   LENGTH: 524
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 479, 484, 486, 487, 496, 501, 505, 506, 509, 511, 513, 514,
```

```
;   LOCATION: 515, 516, 521, 523
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 479, 484, 486, 487, 496, 501, 505, 506, 509, 511, 513, 514,
;   LOCATION: 515, 516, 521, 523
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 479, 484, 486, 487, 496, 501, 505, 506, 509, 511, 513, 514,
;   LOCATION: 515, 516, 521, 523
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 479, 484, 486, 487, 496, 501, 505, 506, 509, 511, 513, 514,
;   LOCATION: 515, 516, 521, 523
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 479, 484, 486, 487, 496, 501, 505, 506, 509, 511, 513, 514,
;   LOCATION: 515, 516, 521, 523
;   OTHER INFORMATION: n = A,T,C or G
US-10-012-697-280

  Query Match             25.6%;  Score 27.4;  DB 15;  Length 524;
  Best Local Similarity   24.3%;  Pred. No. 1.2e+02;
  Matches   25;  Conservative    9;  Mismatches   69;  Indels    0;  Gaps    0;

Qy          2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCU 61
              |||||      : | |      :    :      ||||  :| |||| :     |:
Db        232 GAAAACTCTTGTTAGACATGATGTTTCCTATACAGAAGAAAAGATACAAACTTGATTGCT 291

Qy         62 UNAGNNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAG 104
              : ||             ||| |:               : ||
Db        292 TCAGTAAAAACACTTTTAGAGATTGTGTTTCCTTTTTTGTGAG 334


RESULT 2
US-10-012-697-493
; Sequence 493, Application US/10012697
; Publication No. US20030215803A1
; GENERAL INFORMATION:
;  APPLICANT: Escobedo, Jaime
;  APPLICANT:  Garcia, Pablo Dominguez
;  APPLICANT:  Kassam, Altaf
;  APPLICANT:  Lamson, George
;  APPLICANT:  Scott, Beth
;  APPLICANT:  Drmanac, Radoje
;  APPLICANT:  Crkvenjakov, Radomir
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Drmanac, Snezana
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Leshkowitz, Dena
;  APPLICANT:  Kita, David
;  APPLICANT:  Garcia, Veronica
;  APPLICANT:  Jones, Lee William
;  APPLICANT:  Stache-Crain, Birgit
;  TITLE OF INVENTION: HUMAN GENES AND GENE EXPRESSION PRODUCTS
;  TITLE OF INVENTION:  ISOLATED FROM HUMAN PROSTATE
;  FILE REFERENCE: 2300-16252
;  CURRENT APPLICATION NUMBER: US/10/012,697
;  CURRENT FILING DATE:  2003-01-21
;  PRIOR APPLICATION NUMBER: 60/254,648
;  PRIOR FILING DATE: 2000-12-07
;  PRIOR APPLICATION NUMBER: 60/275,668
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 1568
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 493
;   LENGTH: 761
;   TYPE: DNA
,   ORGANISM: Homo sapiens
,   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 534, 596, 615, 633, 641, 648, 656, 710, 729, 736, 744, 748,
;   LOCATION: 749, 751, 752
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 534, 596, 615, 633, 641, 648, 656, 710, 729, 736, 744, 748,
;   LOCATION: 749, 751, 752
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 534, 596, 615, 633, 641, 648, 656, 710, 729, 736, 744, 748,
;   LOCATION: 749, 751, 752
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 534, 596, 615, 633, 641, 648, 656, 710, 729, 736, 744, 748,
;   LOCATION: 749, 751, 752
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 534, 596, 615, 633, 641, 648, 656, 710, 729, 736, 744, 748,
;   LOCATION: 749, 751, 752
;   OTHER INFORMATION: n = A,T,C or G
US-10-012-697-493

  Query Match             25.6%;  Score 27.4;  DB 15;  Length 761;
  Best Local Similarity   24.3%;  Pred. No. 1.4e+02;
  Matches   25;  Conservative    9;  Mismatches   69;  Indels    0;  Gaps    0;

Qy          2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCU 61
              |||||      : | |      :    :      ||||  :| |||| :     |:
Db        237 GAAAACTCTTGTTAGACATGATGTTTCCTATACAGAAGAAAAGATACAAACTTGATTGCT 296

Qy         62 UNAGNNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAG 104
              : ||             ||| |:               : ||
Db        297 TCAGTAAAAACACTTTTAGAGATTGTGTTTCCTTTTTTGTGAG 339


RESULT 3
US-10-012-697-418
; GENERAL INFORMATION:
;  APPLICANT: Escobedo, Jaime
;  APPLICANT:  Garcia, Pablo Dominguez
;  APPLICANT:  Kassam, Altaf
;  APPLICANT:  Lamson, George
;  APPLICANT:  Scott, Beth
;  APPLICANT:  Drmanac, Radoje
;  APPLICANT:  Crkvenjakov, Radomir
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Drmanac, Snezana
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Leshkowitz, Dena
;  APPLICANT:  Kita, David
;  APPLICANT:  Garcia, Veronica
;  APPLICANT:  Jones, Lee William
;  APPLICANT:  Stache-Crain, Birgit
;  TITLE OF INVENTION: HUMAN GENES AND GENE EXPRESSION PRODUCTS
;  TITLE OF INVENTION:  ISOLATED FROM HUMAN PROSTATE
;  FILE REFERENCE: 2300-16252
;  CURRENT APPLICATION NUMBER: US/10/012,697
;  CURRENT FILING DATE:  2003-01-21
;  PRIOR APPLICATION NUMBER: 60/254,648
;  PRIOR FILING DATE: 2000-12-07
;  PRIOR APPLICATION NUMBER: 60/275,668
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 1568
;  SOFTWARE: FastSEQ for Windows Version 4.0
;   OTHER INFORMATION: n = A,T,C or G
US-10-012-697-418

  Query Match             25.6%;  Score 27.4;  DB 15;  Length 769;
  Best Local Similarity   24.3%;  Pred. No. 1.4e+02;
```

```
  Matches   25;  Conservative    9;  Mismatches   69;  Indels    0;  Gaps    0;

Qy          2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCU 61
              |||||      : | |         :   :        ||||  :| ||||  :     |:
Db         88 GAAAACTCTTGTTAGACATGATGTTTCCTATACAGAAGAAAAGATACAAACTTGATTGCT 147

Qy         62 UNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAG 104
              : ||             ||| |:              : ||
Db        148 TCAGTAAAAACACTTTTAGAGATTGTGTTTCCTTTTTTGTGAG 190


RESULT 4
US-10-012-697-415
  Sequence 415, Application US/10012697
  Publication No. US20030215803A1
  GENERAL INFORMATION:
   APPLICANT: Escobedo, Jaime
   APPLICANT:  Garcia, Pablo Dominguez
   APPLICANT:  Kassam, Altaf
   APPLICANT:  Lamson, George
   APPLICANT:  Scott, Beth
   APPLICANT:  Drmanac, Radoje
   APPLICANT:  Crkvenjakov, Radomir
   APPLICANT:  Dickson, Mark
   APPLICANT:  Drmanac, Snezana
   APPLICANT:  Labat, Ivan
   APPLICANT:  Leshkowitz, Dena
   APPLICANT:  Kita, David
   APPLICANT:  Garcia, Veronica
   APPLICANT:  Jones, Lee William
   APPLICANT:  Stache-Crain, Birgit
   TITLE OF INVENTION: HUMAN GENES AND GENE EXPRESSION PRODUCTS
   TITLE OF INVENTION:  ISOLATED FROM HUMAN PROSTATE
   FILE REFERENCE: 2300-16252
   CURRENT APPLICATION NUMBER: US/10/012,697
   CURRENT FILING DATE:  2003-01-21
   PRIOR APPLICATION NUMBER: 60/254,648
   PRIOR FILING DATE: 2000-12-07
   PRIOR APPLICATION NUMBER: 60/275,668
   PRIOR FILING DATE: 2001-03-13
   NUMBER OF SEQ ID NOS: 1568
   SOFTWARE: FastSEQ for Windows Version 4.0
  SEQ ID NO 415
    LENGTH: 789
    TYPE: DNA
    ORGANISM: Homo sapiens
    FEATURE:
    NAME/KEY: misc_feature
    LOCATION: 5, 7, 9, 12, 15, 17, 19, 20, 27, 33, 46, 62, 63, 64, 66,
    LOCATION: 67, 74, 76, 509, 613, 645, 651, 655, 666, 678, 715, 718, 729, 742, 772, 77
    OTHER INFORMATION: n = A,T,C or G
    FEATURE:
    NAME/KEY: misc_feature
    LOCATION: 5, 7, 9, 12, 15, 17, 19, 20, 27, 33, 46, 62, 63, 64, 66,
    LOCATION: 67, 74, 76, 509, 613, 645, 651, 655, 666, 678, 715, 718, 729, 742, 772, 77
    OTHER INFORMATION: n = A,T,C or G
    FEATURE:
    NAME/KEY: misc_feature
    LOCATION: 5, 7, 9, 12, 15, 17, 19, 20, 27, 33, 46, 62, 63, 64, 66,
    LOCATION: 67, 74, 76, 509, 613, 645, 651, 655, 666, 678, 715, 718, 729, 742, 772, 77
    OTHER INFORMATION: n = A,T,C or G
    FEATURE:
    NAME/KEY: misc_feature
    LOCATION: 5, 7, 9, 12, 15, 17, 19, 20, 27, 33, 46, 62, 63, 64, 66,
    LOCATION: 67, 74, 76, 509, 613, 645, 651, 655, 666, 678, 715, 718, 729, 742, 772, 77
    OTHER INFORMATION: n = A,T,C or G
    FEATURE:
    NAME/KEY: misc_feature
    LOCATION: 5, 7, 9, 12, 15, 17, 19, 20, 27, 33, 46, 62, 63, 64, 66,
    LOCATION: 67, 74, 76, 509, 613, 645, 651, 655, 666, 678, 715, 718, 729, 742, 772, 77
    OTHER INFORMATION: n = A,T,C or G
US-10-012-697-415
```

```
  Query Match           25.6%;  Score 27.4;  DB 15;  Length 789;
  Best Local Similarity   24.3%;  Pred. No. 1.4e+02;
  Matches   25;  Conservative    9;  Mismatches   69;  Indels    0;  Gaps    0;

Qy          2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCU 61
              |||||      : | |         :   :        ||||  :| ||||  :     |:
Db         97 GAAAACTCTTGTTAGACATGATGTTTCCTATACAGAAGAAAAGATACAAACTTGATTGCT 156

Qy         62 UNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAG 104
              : ||             ||| |:              : ||
Db        157 TCAGTAAAAACACTTTTAGAGATTGTGTTTCCTTTTTTGTGAG 199


RESULT 5
US-10-012-697-1307
; Sequence 1307, Application US/10012697
; Publication No. US20030215803A1
; GENERAL INFORMATION:
;  APPLICANT: Escobedo, Jaime
;  APPLICANT:  Garcia, Pablo Dominguez
;  APPLICANT:  Kassam, Altaf
;  APPLICANT:  Lamson, George
;  APPLICANT:  Scott, Beth
;  APPLICANT:  Drmanac, Radoje
;  APPLICANT:  Crkvenjakov, Radomir
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Drmanac, Snezana
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Leshkowitz, Dena
;  APPLICANT:  Kita, David
;  APPLICANT:  Garcia, Veronica
;  APPLICANT:  Jones, Lee William
;  APPLICANT:  Stache-Crain, Birgit
;  TITLE OF INVENTION: HUMAN GENES AND GENE EXPRESSION PRODUCTS
;  TITLE OF INVENTION:  ISOLATED FROM HUMAN PROSTATE
;  FILE REFERENCE: 2300-16252
;  CURRENT APPLICATION NUMBER: US/10/012,697
;  CURRENT FILING DATE:  2003-01-21
;  PRIOR APPLICATION NUMBER: 60/254,648
;  PRIOR FILING DATE: 2000-12-07
;  PRIOR APPLICATION NUMBER: 60/275,668
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 1568
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1307
;   LENGTH: 929
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 649, 785, 791, 795, 806, 818, 855, 858, 869, 882, 912, 916,
;   LOCATION: 919, 920, 921, 923, 926, 927, 928
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 649, 785, 791, 795, 806, 818, 855, 858, 869, 882, 912, 916,
;   LOCATION: 919, 920, 921, 923, 926, 927, 928
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 649, 785, 791, 795, 806, 818, 855, 858, 869, 882, 912, 916,
;   LOCATION: 919, 920, 921, 923, 926, 927, 928
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 649, 785, 791, 795, 806, 818, 855, 858, 869, 882, 912, 916,
;   LOCATION: 919, 920, 921, 923, 926, 927, 928
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 649, 785, 791, 795, 806, 818, 855, 858, 869, 882, 912, 916,
;   LOCATION: 919, 920, 921, 923, 926, 927, 928
;   OTHER INFORMATION: n = A,T,C or G
```

US-10-012-697-1307

Query Match 25.6%; Score 27.4; DB 15; Length 929;
Best Local Similarity 24.3%; Pred. No. 1.5e+02;
Matches 25; Conservative 9; Mismatches 69; Indels 0; Gaps 0;

```
Qy          2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCU 61
              |||||      : | |         :   :       ||||  :| |||| :     |:
Db        237 GAAAACTCTTGTTAGACATGATGTTTCCTATACAGAAGAAAAGATACAAACTTGATTGCT 296

Qy         62 UNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUAG 104
              : ||             ||| |:             : ||
Db        297 TCAGTAAAAACACTTTTAGAGATTGTGTTTCCTTTTTTGTGAG 339
```

RESULT 6
US-10-311-455-239/c
; Sequence 239, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 239
; LENGTH: 6419
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-239

Query Match 25.0%; Score 26.8; DB 14; Length 6419;
Best Local Similarity 24.0%; Pred. No. 4.4e+02;
Matches 25; Conservative 9; Mismatches 70; Indels 0; Gaps 0;

```
Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              ||||      : |:|      |     ::      |||||    ||| |:     |::
Db       1272 AAAAAATAAATTATAATTAATTTCCCATCTTACAAATAAAAAAACTAAAAATCGAAACTT 1213

Qy         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUAGCA 106
              |              | |||:                 :| ||
Db       1212 AAAAAACAACCTACCCAAAAATCACCCAACAAACTATTCTAACA 1169
```

RESULT 7
US-10-311-455-1166/c
; Sequence 1166, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1166
; LENGTH: 13814
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 9297
; OTHER INFORMATION: n is a or g or c or t
US-10-311-455-1166

Query Match 25.0%; Score 26.8; DB 14; Length 13814;
Best Local Similarity 22.1%; Pred. No. 5.9e+02;
Matches 23; Conservative 11; Mismatches 70; Indels 0; Gaps 0;

```
Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              ||||:      : | |:          :::       |||||  | |  | :     |:
Db       2230 AAAATTAATTTTAAATTTTCTAAAACACTTTTTAACAAAAAAAAAAATTCTTAAACACTA 2171

Qy         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUAGCA 106
              |              | |||:                 ::| ||
Db       2170 AAAACAAAATTAAATTAAAAATATTCAATACTCTTCCATTATCA 2127
```

RESULT 8
US-10-312-841-2/c
; Sequence 2, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MH(
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 2
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (379615)
US-10-312-841-2

Query Match 25.0%; Score 26.8; DB 14; Length 3673778;
Best Local Similarity 22.1%; Pred. No. 9.4e+02;
Matches 23; Conservative 11; Mismatches 70; Indels 0; Gaps 0;

```
Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              ||||:      : |:|:          :   :::       | |||     |||||:
Db    2995750 AAAATTATCTTCATATAAAAACCATAAATTTATCTATATAAATCTCAAACATATCATTAA 2995691

Qy         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUAGCA 106
              |              | |||:                 : | |||
Db    2995690 AATATAACATACCCAAAAAAATTTAATTAATATATATTTAAACA 2995647
```

RESULT 9
US-10-027-632-323248/c
; Sequence 323248, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632

```
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 323248
;   LENGTH: 487
;   TYPE: DNA
;   ORGANISM: Human
JS-10-027-632-323248

  Query Match             24.9%;  Score 26.6;  DB 15;  Length 487;
  Best Local Similarity   22.2%;  Pred. No. 2e+02;
  Matches   22;  Conservative   10;  Mismatches   67;  Indels    0;  Gaps    0;

Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              ||||         |: :|       |:   :::       ||||| :| ||| |       | :
Db        378 AAAAGAAAAAAAATTTGCAACTTCTTTCTTTTGCAAGAAAAAATAAAAAGACAACCACAT 319

Qy         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNU 101
              ||           ||:             :
Db        318 AAGTTACCAGAATGTGCATAATCATTACTTCATGATCTT 280


RESULT 10
JS-10-027-632-324208/c
; Sequence 324208, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 324208
;   LENGTH: 488
;   TYPE: DNA
;   ORGANISM: Human
JS-10-027-632-324208

  Query Match             24.9%;  Score 26.6;  DB 15;  Length 488;
  Best Local Similarity   22.2%;  Pred. No. 2e+02;
  Matches   22;  Conservative   10;  Mismatches   67;  Indels    0;  Gaps    0;

Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              ||||         |: :|       |:   :::       ||||| :| ||| |       | :
Db        378 AAAAGAAAAAAAATTTGCAACTTCTTTCTTTTGCAAGAAAAAATAAAAAGACAACCACAT 319

Qy         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNU 101
              ||           ||:             :
Db        318 AAGTTACCAGAATGTGCATAATCATTACTTCATGATCTT 280


RESULT 11
US-10-027-632-8286
; Sequence 8286, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8286
;   LENGTH: 634
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-8286

  Query Match             24.9%;  Score 26.6;  DB 15;  Length 634;
  Best Local Similarity   23.4%;  Pred. No. 2.2e+02;
  Matches   25;  Conservative   10;  Mismatches   72;  Indels    0;  Gaps    0;

Qy          1 GGAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNNAAAAANUANAAACAUNNNNNC 60
              |||| |      :  : :|       |    ::       ||||   | | |||:
Db        421 GGAAGAGACCATTCTTTTGCATACCCCATTTTCTGGATAAAAAGAAAAAGACATTCCTGT 480

Qy         61 UUNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUAGCAG 107
              ::  |              |||||:                : | | |
Db        481 TTTTGATCTTTTAATTCAAGAAATGGGGCAAAAACCAACCTAATCTG 527


RESULT 12
US-10-311-455-1858
; Sequence 1858, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
;  APPLICANT: OLEK, Alexander
;  APPLICANT:  PIEPENBROCK, Christian
;  APPLICANT:  BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
;  TITLE OF INVENTION:  cytosine methylation
;  FILE REFERENCE: 5013.1014
;  CURRENT APPLICATION NUMBER: US/10/311,455
;  CURRENT FILING DATE:  2002-12-16
;  PRIOR APPLICATION NUMBER: PCT/EP01/07537
;  PRIOR FILING DATE: 2001-07-02
;  PRIOR APPLICATION NUMBER: DE 10032529.7
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: DE 10043826.1
```

```
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1858
;   LENGTH: 10034
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1858

  Query Match             24.9%;  Score 26.6;  DB 14;  Length 10034;
  Best Local Similarity   17.2%;  Pred. No. 6e+02;
  Matches   17;  Conservative   15;  Mismatches   67;  Indels    0;  Gaps    0;

Qy          4 AAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUUN 63
              |||:     : |:|:|      :   :::      ||||  :| ||  |:      ::
Db       2863 AAATTTGATTAATATGGTGAAATTTTATTTTTATTAAAAATATAAAATTATTGGGTATTT 2922

Qy         64 AGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUU 102
              |              || :                ::
Db       2923 TGGTAGCGTATGTTTGTAATTTTAGTTGTGTAGGAGGTT 2961


RESULT 13
US-09-814-353-17808/c
; Sequence 17808, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
;  APPLICANT: Lee, John
;  APPLICANT:  Thompson, Pamela
;  APPLICANT:  Lillie, James
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;  TITLE OF INVENTION:  IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF OVARIAN CANCER
;  FILE REFERENCE: MRI-006B
;  CURRENT APPLICATION NUMBER: US/09/814,353
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: US 60/191,031
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: US 60/207,124
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: US 60/211,940
;  PRIOR FILING DATE: 2000-06-15
;  PRIOR APPLICATION NUMBER: US 60/216,820
;  PRIOR FILING DATE: 2000-07-07
;  PRIOR APPLICATION NUMBER: US 60/220,661
;  PRIOR FILING DATE: 2000-07-25
;  PRIOR APPLICATION NUMBER: US 60/257,672
;  PRIOR FILING DATE: 2000-12-21
;  NUMBER OF SEQ ID NOS: 22037
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17808
;   LENGTH: 390
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-814-353-17808

  Query Match             24.5%;  Score 26.2;  DB 10;  Length 390;
  Best Local Similarity   23.5%;  Pred. No. 2.4e+02;
  Matches   19;  Conservative   12;  Mismatches   50;  Indels    0;  Gaps    0;

Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              ||||:     : |: :      :   :::      |||||  :| |||  |:      ::
Db        309 AAAATAATTTTTATTTTTTTCTTTTTTTTTTTTTTTTAAAAATTAAAAAAATTAATTTTT 250

Qy         63 NAGNNNNNNNNNNNNNAGAAA 83
              |  |||
Db        249 TTTTTTTTTTAAAAAAAAAA 229


RESULT 14
US-10-312-841-1/c
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
;  APPLICANT: Epigenomics AG
;  TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MH
;  FILE REFERENCE: E01/1208/WO
;  CURRENT APPLICATION NUMBER: US/10/312,841
;  CURRENT FILING DATE:  2002-12-30
;  NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
;   LENGTH: 3673778
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (3294164)
US-10-312-841-1

  Query Match             24.5%;  Score 26.2;  DB 14;  Length 3673778;
  Best Local Similarity   22.9%;  Pred. No. 9.4e+02;
  Matches   24;  Conservative   10;  Mismatches   71;  Indels    0;  Gaps    0;

Qy          2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCU 61
              ||||      : |:|:       |:   :::      |||||  |  || |       :
Db    1459596 GAAACCTACCGTAATATTAAATATCTAAATTTTAATTAAAAAAAAAACAAAAAACCTAAT 1459537

Qy         62 UNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUAGCA 106
              |                        | |||                ::| ||
Db    1459536 ATACAAATTCCATAATAAAAAAACCCAAAACGAATATCTTATCA 1459492


RESULT 15
US-10-424-599-85840
; Sequence 85840, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 85840
;   LENGTH: 511
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_48527C.1
US-10-424-599-85840

  Query Match             24.3%;  Score 26;  DB 12;  Length 511;
  Best Local Similarity   24.0%;  Pred. No. 3e+02;
  Matches   24;  Conservative    8;  Mismatches   68;  Indels    0;  Gaps    0;

Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              ||||      | |:|       |:   :::      |||||  | ||||       |
Db        270 AAAAGGGAACCAAAATGTCCCCACTATTTTTAATAAAAAAAAAAAAAAACCAAATATCGG 329

Qy         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUU 102
              |              || ||:                ::
Db        330 AGGGGGGGGGGTTAAAAGGAATAAAAGTTAACGTCCCCTT 369


Search completed: March  8, 2004, 06:22:39
Job time : 1232 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

)M nucleic - nucleic search, using sw model

Run on: March 8, 2004, 04:24:32 ; Search time 2381 Seconds
(without alignments)
1341.980 Million cell updates/sec

Title: US-09-646-651C-2
Perfect score: 107
Sequence: 1 ggaaaaunnnnnunauaugn.....nnnnnnnnnnnnnnuuagcag 107

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

esult No.		Score	% Query Match	Length	DB	ID	Description
c	1	29.6	27.7	679	28	BZ043484	BZ043484 lkg57f11.
	2	29	27.1	453	10	BG059751	BG059751 naf53b08.
c	3	28.8	26.9	248	13	BX684795	BX684795 BX684795
	4	28.8	26.9	685	28	BH488315	BH488315 BOGWD90TF
c	5	28.8	26.9	794	28	BH489963	BH489963 BOGKA57TR
c	6	28.6	26.7	1125	29	CNS057N6	AL324843 Tetraodon
	7	28	26.2	367	28	AQ135026	AQ135026 HS_3053_A
	8	28	26.2	394	28	AQ205663	AQ205663 HS_3236_B
	9	28	26.2	394	28	AQ268612	AQ268612 RPCI11-71
	10	28	26.2	406	28	AQ211558	AQ211558 HS_3235_B
	11	27.8	26.0	539	9	AL915410	AL915410 AL915410
	12	27.8	26.0	539	9	AL915411	AL915411 AL915411
	13	27.6	25.8	618	28	BH719067	BH719067 BOHXL74TF
	14	27.6	25.8	807	28	BH654909	BH654909 BOHTZ17TF
	15	27.4	25.6	204	14	CB050688	CB050688 NISC_gj19
	16	27.4	25.6	508	14	CB050532	CB050532 NISC_gj18
c	17	27.4	25.6	544	14	CB050531	CB050531 NISC_gj18
c	18	27.4	25.6	704	9	AL720964	AL720964 AL720964
	19	27	25.2	234	28	CC397779	CC397779 PUHFZ58TD
	20	27	25.2	500	9	AU087590	AU087590 AU087590
	21	27	25.2	714	28	BH970555	BH970555 odj11g05.
	22	27	25.2	730	28	BZ032632	BZ032632 oei22f07.
	23	27	25.2	779	28	CC098521	CC098521 CSU-K34.1
	24	27	25.2	810	28	BH480090	BH480090 BOGCF72TF
	25	27	25.2	811	28	BH250315	BH250315 BOGAI90TF
c	26	27	25.2	860	12	BG976823	BG976823 602845725
c	27	27	25.2	942	14	CB182382	CB182382 AGENCOURT
c	28	27	25.2	1101	29	CNS0177G	AL107686 Drosophil
c	29	26.8	25.0	333	14	CF544518	CF544518 Hd_mx17_6
	30	26.8	25.0	463	28	AQ466253	AQ466253 HS_5157_A
	31	26.8	25.0	540	28	AZ067096	AZ067096 RPCI-23-4
	32	26.8	25.0	653	29	CE241947	CE241947 tigr-gss-
	33	26.8	25.0	686	29	CE846494	CE846494 tigr-gss-
c	34	26.8	25.0	963	29	CNS074X6	AL429376 clone BA0
c	35	26.6	24.9	429	28	AQ513768	AQ513768 HS_5152_A
c	36	26.6	24.9	458	28	AQ517382	AQ517382 HS_5135_B
c	37	26.6	24.9	487	28	BZ173324	BZ173324 CH230-321
c	38	26.6	24.9	719	29	AG008606	AG008606 Homo sapi
c	39	26.6	24.9	739	14	CF259384	CF259384 388 Trich
	40	26.6	24.9	781	28	CC077740	CC077740 CSU-K33r.
	41	26.6	24.9	841	28	BZ262911	BZ262911 CH230-374
	42	26.6	24.9	857	29	CNS010S3	AL099357 Drosophil
	43	26.6	24.9	859	28	CC198967	CC198967 ZMMBBc032
c	44	26.4	24.7	182	28	BH911389	BH911389 SALK_0682
c	45	26.4	24.7	418	12	BG941688	BG941688 ax16b02.x

ALIGNMENTS

RESULT 1
BZ043484/c
LOCUS BZ043484 679 bp DNA linear GSS 09-OCT-2002
DEFINITION lkg57f11.b1 B.oleracea002 Brassica oleracea genomic, genomic survey sequence.
ACCESSION BZ043484
VERSION BZ043484.1 GI:23635947
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 679)
AUTHORS Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T., Nash,W., Rabinowicz,P.D. and Wilson,R.K.
TITLE Whole genome shotgun reads from Brassica oleracea
JOURNAL Unpublished (2002)
COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Plate: lkg57 row: f column: 11
Seq primer: -21UPpOT forward
Class: shotgun
High quality sequence start: 17
High quality sequence stop: 551.

```
FEATURES             Location/Qualifiers
     source          1. .679
                     /organism="Brassica oleracea"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3712"
                     /clone_lib="B.oleracea002"
                     /note="Vector: pOTw13; Whole genome shotgun library from
                     flowering buds. DNA was purified from a crude nuclear
                     prep using Brassica oleracea TO1000DH3 buds provided by
                     Thosmas Osborn at the University of Wisconsin. Genomic
                     DNA was provided by Pablo Rabinowicz (CSHL) and the
                     shotgun library prepared at Washington University Genome
                     Sequencing Center."
ORIGIN

  Query Match              27.7%;  Score 29.6;  DB 28;  Length 679;
  Best Local Similarity    23.5%;  Pred. No. 1.1e+03;
  Matches   24;  Conservative   11;  Mismatches   67;  Indels    0;  Gaps    0;

Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              ||||       |:|:       |:   :::      |||||  :|  ||  |:     |:
Db        333 AAAAATACATAGATATATTACCTCTAGATTTGTTTAAAAAAATTATTAAAATCTTTACTA 274

Qy         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUAG 104
                |                | |||:          :||
Db        273 AATCGTTTACAACCTCAAAAATCTCAGAATCGGCGGAAGTAG 232


RESULT 2
BG059751
LOCUS       BG059751                453 bp    mRNA    linear   EST 25-JAN-2001
DEFINITION  naf53b08.x1 NCI_CGAP_Brn65 Homo sapiens cDNA clone IMAGE:4147551 3'
            similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION   BG059751
VERSION     BG059751.1  GI:12527546
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 453)
  AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Neurological
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 421.
FEATURES             Location/Qualifiers
     source          1. .453
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4147551"
                     /tissue_type="glioblastoma without EGFR amplification"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NCI_CGAP_Brn65"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.77 kb. Constructed by Life
                     Technologies."
ORIGIN

  Query Match              27.1%;  Score 29;  DB 10;  Length 453;
  Best Local Similarity    23.3%;  Pred. No. 1.7e+03;
  Matches   24;  Conservative   11;  Mismatches   68;  Indels    0;  Gaps    0;

Qy          1 GGAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNC 60
              |||  ||:       : |  :       |:   :::      |||  | :  |||||:
Db        189 GGAGAATGAAATTAACTTTTTCTGTCTGGATTTAGGAGAAAACAATGGAAACATTTTAAA 248

Qy         61 UUNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUA 103
                ||              |||||                ::|
Db        249 ACAAGTTTTATATCTGAGAGAAAAAAAAGGAATGATTCCATTA 291


RESULT 3
BX684795/c
LOCUS       BX684795                248 bp    mRNA    linear   EST 14-NOV-2003
DEFINITION  BX684795 XGC-neurula Silurana tropicalis cDNA clone TNeu075o20 3',
            mRNA sequence.
ACCESSION   BX684795
VERSION     BX684795.1  GI:38333915
KEYWORDS    EST.
SOURCE      Silurana tropicalis (western clawed frog)
  ORGANISM  Silurana tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Silurana.
REFERENCE   1  (bases 1 to 248)
  AUTHORS   Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
  TITLE     Sanger Xenopus tropicalis EST project 2001 (11_2003)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Croning MDR
            Sanger Institute
            Hinxton, Cambridgeshire, CB10 1SA, UK
            Email: trop@sanger.ac.uk
            Sanger Xenopus tropicalis EST project 2001
            TROPICALIS_SEQUENCE_ID: TNeu075o20.q1kT7
            Sequencing primer: T7
            This sequence is from a Xenopus Gene Collection (XGC) library
            constructed by Aaron M. Zorn.
            cDNA was oligo dT primed from 5ug of poly A+ RNA from neurula.
            EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
            5' end and NotI at the 3' end.
            Vector: pCS107; Site_1: EcoRI; Site_2: NotI
            Host: Escherichia coli DH10B.
FEATURES             Location/Qualifiers
     source          1. .248
                     /organism="Silurana tropicalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8364"
                     /clone="TNeu075o20"
                     /dev_stage="neurula"
                     /lab_host="Escherichia coli DH10B"
                     /clone_lib="XGC-neurula"
                     /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
                     was oligo dT primed from 5ug of poly A+ RNA from neurula.
                     EcoRI-NotI cut cDNA was then ligated into pCS107 with
                     EcoRI at the 5' end and NotI at the 3' end."
ORIGIN

  Query Match              26.9%;  Score 28.8;  DB 13;  Length 248;
  Best Local Similarity    31.7%;  Pred. No. 2.3e+03;
  Matches   26;  Conservative    8;  Mismatches   48;  Indels    0;  Gaps    0;

Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              ||||:      : | | |       |:   :::      |||||  | |||  |:     |
Db        110 AAAATGATGGTAAGAGGTTTACCCTTTTTTTCAAAAAAAAAAAAAAAAAAAATTTCAACGG 51

Qy         63 NAGNNNNNNNNNNNNNAGAAAU 84
                |||             | |||:
Db         50 GAGNTAAAAAAAAAAAAAAAT 29


RESULT 4
BH488315
```

```
LOCUS       BH488315                685 bp    DNA     linear   GSS 13-DEC-2001
DEFINITION  BOGWD90TF BOGW Brassica oleracea genomic clone BOGWD90, genomic
            survey sequence.
ACCESSION   BH488315
VERSION     BH488315.1  GI:17696419
KEYWORDS    GSS.
SOURCE      Brassica oleracea
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 685)
  AUTHORS   Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
  TITLE     Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .685
                     /organism="Brassica oleracea"
                     /mol_type="genomic DNA"
                     /strain="TO1000DH3"
                     /db_xref="taxon:3712"
                     /clone="BOGWD90"
                     /clone_lib="BOGW"
                     /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
                     genomic DNA inserted into pHOS1 using BstXI linkers"
ORIGIN

  Query Match             26.9%;  Score 28.8;  DB 28;  Length 685;
  Best Local Similarity   24.5%;  Pred. No. 1.6e+03;
  Matches   26;  Conservative   10;  Mismatches   70;  Indels    0;  Gaps    0;

Qy          1 GGAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNC 60
              ||||||:     : |:| |       :   ::        ||||  :  | |||      |
Db        463 GGAAAATATATATTATAAGAGAATTTTTTTTTATAGAATGAAAATTTCACACAAATTTTC 522

Qy         61 UUNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUAGCA 106
              ::                |||||:                 | ||
Db        523 TTCCATCAACAATTACTCAGAAATAACCAAAGCTGCTCTCGCATCA 568


RESULT 5
BH489963/c
LOCUS       BH489963                794 bp    DNA     linear   GSS 13-DEC-2001
DEFINITION  BOGKA57TR BOGK Brassica oleracea genomic clone BOGKA57, genomic
            survey sequence.
ACCESSION   BH489963
VERSION     BH489963.1  GI:17698067
KEYWORDS    GSS.
SOURCE      Brassica oleracea
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 794)
  AUTHORS   Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
  TITLE     Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL   Unpublished (2001)
COMMENT     Other_GSSs: BOGKA57TF
            Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .794
                     /organism="Brassica oleracea"
                     /mol_type="genomic DNA"
                     /strain="TO1000DH3"
                     /db_xref="taxon:3712"
                     /clone="BOGKA57"
                     /clone_lib="BOGK"
                     /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
                     genomic DNA inserted into pHOS1 using BstXI linkers"
ORIGIN

  Query Match             26.9%;  Score 28.8;  DB 28;  Length 794;
  Best Local Similarity   24.5%;  Pred. No. 1.5e+03;
  Matches   26;  Conservative   10;  Mismatches   70;  Indels    0;  Gaps    0;

Qy          1 GGAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNC 60
              ||||||:     : |:| |       :   ::        ||||  :  | |||      |
Db        128 GGAAAATATATATTATAAGAGAATTTTTTTTTATAGAATGAAAATTTCACACAAATTTTC 69

Qy         61 UUNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUAGCA 106
              ::                |||||:                 | ||
Db         68 TTCCATCAACAATTACTCAGAAATAACCAAAGCTGCTCTCGCATCA 23


RESULT 6
CNS057N6/c
LOCUS       CNS057N6                1125 bp    DNA     linear   GSS 26-JUL-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence T7 end of clone
            010K04 of library A from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL324843
VERSION     AL324843.1  GI:9557725
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetradontoidea; Tetraodontidae; Tetraodon.
REFERENCE   1
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Estimate of human gene number provided by genome-wide analysis
            using Tetraodon nigroviridis DNA sequence
  JOURNAL   Nat. Genet. 25 (2), 235-238 (2000)
  MEDLINE   20296633
   PUBMED   10835645
REFERENCE   2
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
            Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
            Saurin,W., Bernot,A. and Weissenbach,J.
  TITLE     Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Genome Res. 10 (7), 939-949 (2000)
  MEDLINE   20359837
   PUBMED   10899143
REFERENCE   3  (bases 1 to 1125)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1. .1125
```

```
                     /organism="Tetraodon nigroviridis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:99883"
                     /clone="010K04"
                     /clone_lib="A"
                     /note="Genoscope sequence ID : C0AA010BF02C3~end : T7"
ORIGIN

  Query Match              26.7%;  Score 28.6;  DB 29;  Length 1125;
  Best Local Similarity    21.8%;  Pred. No. 1.5e+03;
  Matches   22;  Conservative   12;  Mismatches   67;  Indels    0;  Gaps    0;

Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              ||||:      |:|:       :  :::       ||||  | |||||:     |::
Db        247 AAAATAACTAAAATATTTTTAACTTGCATTTATATATCAAAAGCATAAACATTAACTCTT 188

Qy         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUA 103
                          | |||             ::|
Db        187 AGCTTTTATCAGATGAACAAAGTGTATTTGAGGTCTGTTTA 147


RESULT 7
AQ135026
LOCUS       AQ135026                367 bp    DNA     linear   GSS 24-SEP-1998
DEFINITION  HS_3053_A2_A11_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3053 Col=22 Row=A, genomic survey
            sequence.
ACCESSION   AQ135026
VERSION     AQ135026.1  GI:3526392
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 367)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 3053  row: A  column: 22
            Class: BAC ends
            High quality sequence stop: 367.
FEATURES             Location/Qualifiers
     source          1. .367
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=3053 Col=22 Row=A"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
ORIGIN

  Query Match              26.2%;  Score 28;  DB 28;  Length 367;
  Best Local Similarity    23.5%;  Pred. No. 3.1e+03;
  Matches   24;  Conservative   10;  Mismatches   68;  Indels    0;  Gaps    0;

Qy          2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCU 61
              |||||:      : | |:             ::       || || :|  ||||:     |
Db        233 GAAAATCTCTTTAAAATATACTTGACTGCTTAAAGCAAAAGAATTAGTAACATGACATCC 292

Qy         62 UNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUA 103
              : ||             |||| :                 :|
Db        293 TAAGGTTTATAATGTGTAGAAGTAAAGTGTATTATAACCATA 334


RESULT 8
AQ205663
LOCUS       AQ205663                394 bp    DNA     linear   GSS 17-SEP-1998
DEFINITION  HS_3236_B2_A03_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3236 Col=6 Row=B, genomic survey
            sequence.
ACCESSION   AQ205663
VERSION     AQ205663.1  GI:3616233
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 394)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 3236  row: B  column: 6
            Class: BAC ends
            High quality sequence stop: 394.
FEATURES             Location/Qualifiers
     source          1. .394
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=3236 Col=6 Row=B"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
ORIGIN

  Query Match              26.2%;  Score 28;  DB 28;  Length 394;
  Best Local Similarity    23.5%;  Pred. No. 3e+03;
  Matches   24;  Conservative   10;  Mismatches   68;  Indels    0;  Gaps    0;

Qy          2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCU 61
              |||||:      : | |:             ::       || || :|  ||||:     |
Db        233 GAAAATCTCTTTAAAATATACTTGACTGCTTAAAGCAAAAGAATTAGTAACATGACATCC 292

Qy         62 UNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUA 103
              : ||             |||| :                 :|
Db        293 TAAGGTTTATAATGTGTAGAAGTAAAGTGTATTATAACCATA 334


RESULT 9
AQ268612
LOCUS       AQ268612                394 bp    DNA     linear   GSS 27-APR-1999
DEFINITION  RPCI11-71J23.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-71J23,
            genomic survey sequence.
ACCESSION   AQ268612
VERSION     AQ268612.1  GI:3796216
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
EFERENCE    1  (bases 1 to 394)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
  TITLE     Use of human BAC End Sequences for Sequence-Ready Map Building
  JOURNAL   Unpublished (1998)
OMMENT      Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: SP6
            Class: BAC ends.
EATURES                 Location/Qualifiers
     source             1. .394
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="GDB:7527118"
                        /db_xref="taxon:9606"
                        /clone="RPCI-11-71J23"
                        /sex="Male"
                        /cell_type="Lymphocytes"
                        /clone_lib="RPCI-11"
                        /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                        RPCI11 Human Male BAC Library"
RIGIN

  Query Match             26.2%;  Score 28;  DB 28;  Length 394;
  Best Local Similarity   23.5%;  Pred. No. 3e+03;
  Matches   24;  Conservative   10;  Mismatches   68;  Indels    0;  Gaps    0;

y           2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCU 61
              |||||:     : | |:            ::       || || :|  ||||:      |
b         201 GAAAATCTCTTTAAAATATACTTGACTGCTTAAAGCAAAAGAATTAATAACATGACATCC 260

y          62 UNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUA 103
             : ||              |||| :                 :|
b         261 TAAGGTTTATAATGTGTAGAAGTAAAGTGTATTATAACCATA 302


ESULT 10
Q211558
OCUS        AQ211558                406 bp    DNA     linear   GSS 18-SEP-1998
EFINITION   HS_3235_B2_E01_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3235 Col=2 Row=J, genomic survey
            sequence.
CCESSION    AQ211558
ERSION      AQ211558.1  GI:3622759
EYWORDS     GSS.
OURCE       Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
EFERENCE    1  (bases 1 to 406)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
    PUBMED  10449764
OMMENT      Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 3235  row: J  column: 2
            Class: BAC ends
            High quality sequence stop: 406.
FEATURES                Location/Qualifiers
     source             1. .406
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /clone="Plate=3235 Col=2 Row=J"
                        /sex="male"
                        /clone_lib="CIT Approved Human Genomic Sperm Library D"
                        /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                        E-Coli DH10B"
ORIGIN

  Query Match             26.2%;  Score 28;  DB 28;  Length 406;
  Best Local Similarity   23.5%;  Pred. No. 3e+03;
  Matches   24;  Conservative   10;  Mismatches   68;  Indels    0;  Gaps    0;

Qy          2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCU 61
              |||||:     : | |:            ::       || || :|  ||||:      |
Db        233 GAAAATCTCTTTAAAATATACTTGACTGCTTAAAGCAAAAGAATTAGTAACATGACATCC 292

Qy         62 UNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUA 103
              : ||              |||| :                 :|
Db        293 TAAGGTTTATAATGTGTAGAAGTAAAGTGTATTATAACCATA 334


RESULT 11
AL915410
LOCUS       AL915410                539 bp    mRNA    linear   EST 18-SEP-2002
DEFINITION  AL915410 PJR-Z1+Z2 Danio rerio cDNA clone 193-E07-1, mRNA sequence.
ACCESSION   AL915410
VERSION     AL915410.1  GI:23180708
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 539)
  AUTHORS   Lo,J., Lee,S., Xu,M., Liu,F., Ruan,H., Eun,A., He,Y., Ma,W.,
            Wang,W., Wen,Z. and Peng,J.
  TITLE     15000 unique zebrafish EST clusters and their future use in
            microarray for profiling gene expression patterns during
            embryogenesis
  JOURNAL   Genome Res. 13 (3), 455-456 (2003)
  MEDLINE   22505427
    PUBMED  12618376
COMMENT     Contact: Peng J
            Lab of Functional Genomics
            Institute of Molecular and Cell Biology
            30 Medical Drive, Singapore, 117609, Singapore
            Email: pengjr@imcb.a-star.edu.sg
            Clone requests: pengjr@imcb.a-star.edu.sg.
FEATURES                Location/Qualifiers
     source             1. .539
                        /organism="Danio rerio"
                        /mol_type="mRNA"
                        /strain="local wildtype"
                        /db_xref="taxon:7955"
                        /clone="193-E07-1"
                        /tissue_type="whole embryo or fish"
                        /dev_stage="mixed stages"
                        /clone_lib="PJR-Z1+Z2"
ORIGIN
```

```
  Query Match             26.0%;  Score 27.8;  DB 9;  Length 539;
  Best Local Similarity   25.7%;  Pred. No. 2.9e+03;
  Matches   27;  Conservative    8;  Mismatches   70;  Indels    0;  Gaps    0;

Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              ||||         | |:|      |:  : :      |||   :| ||| |:      ::
Db          7 AAAACTGGAACAAAATGAAAACACTTAGTCTAAACTAAAATCGTATAAAGATTTGGGATT 66

Qy         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAGCAG 107
               ||          | |||                  |||||
Db         67 GAGCATACAGAGGGCAAAAAAAAAAAAAAAAAAAGTGATAAAGCAG 111


RESULT 12
AL915411
LOCUS       AL915411                539 bp    mRNA    linear   EST 18-SEP-2002
DEFINITION  AL915411 PJR-Z1+Z2 Danio rerio cDNA clone 192-E07-1, mRNA sequence.
ACCESSION   AL915411
VERSION     AL915411.1  GI:23180709
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 539)
  AUTHORS   Lo,J., Lee,S., Xu,M., Liu,F., Ruan,H., Eun,A., He,Y., Ma,W.,
            Wang,W., Wen,Z. and Peng,J.
  TITLE     15000 unique zebrafish EST clusters and their future use in
            microarray for profiling gene expression patterns during
            embryogenesis
  JOURNAL   Genome Res. 13 (3), 455-456 (2003)
  MEDLINE   22505427
   PUBMED   12618376
COMMENT     Contact: Peng J
            Lab of Functional Genomics
            Institute of Molecular and Cell Biology
            30 Medical Drive, Singapore, 117609, Singapore
            Email: pengjr@imcb.a-star.edu.sg
            Clone requests: pengjr@imcb.a-star.edu.sg.
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /strain="local wildtype"
                     /db_xref="taxon:7955"
                     /clone="192-E07-1"
                     /tissue_type="whole embryo or fish"
                     /dev_stage="mixed stages"
                     /clone_lib="PJR-Z1+Z2"
ORIGIN

  Query Match             26.0%;  Score 27.8;  DB 9;  Length 539;
  Best Local Similarity   25.7%;  Pred. No. 2.9e+03;
  Matches   27;  Conservative    8;  Mismatches   70;  Indels    0;  Gaps    0;

Qy          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNCUU 62
              ||||         | |:|      |:  : :      |||   :| ||| |:      ::
Db          7 AAAACTGGAACAAAATGAAAACACTTAGTCTAAACTAAAATCGTATAAAGATTTGGGATT 66

Qy         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNUUAGCAG 107
               ||          | |||                  |||||
Db         67 GAGCATACAGAGGGCAAAAAAAAAAAAAAAAAAAGTGATAAAGCAG 111


RESULT 13
BH719067
LOCUS       BH719067                618 bp    DNA     linear   GSS 20-FEB-2002
DEFINITION  BOHXL74TF BO_2_3_KB Brassica oleracea genomic clone BOHXL74,
            genomic survey sequence.
ACCESSION   BH719067
VERSION     BH719067.1  GI:18818288
KEYWORDS    GSS.
SOURCE      Brassica oleracea
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 618)
  AUTHORS   Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
  TITLE     Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL   Unpublished (2001)
COMMENT     Other_GSSs: BOHXL74TR
            Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .618
                     /organism="Brassica oleracea"
                     /mol_type="genomic DNA"
                     /strain="TO1000DH3"
                     /db_xref="taxon:3712"
                     /clone="BOHXL74"
                     /clone_lib="BO_2_3_KB"
                     /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
                     genomic DNA inserted into pHOS1 using BstXI linkers"
ORIGIN

  Query Match             25.8%;  Score 27.6;  DB 28;  Length 618;
  Best Local Similarity   27.4%;  Pred. No. 3.1e+03;
  Matches   23;  Conservative   10;  Mismatches   51;  Indels    0;  Gaps    0;

Qy          1 GGAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAAACAUNNNNNC 60
              ||||||:     : |:| |       :   ::        |||| : | |||      |
Db        504 GGAAAATATATATTATAAGAGAATTTTTTTTTATAGAATGAAAATTTCACACAAATTTTC 563

Qy         61 UUNAGNNNNNNNNNNNNNAGAAAU 84
              ::              |||||:
Db        564 TTTCATCAACAATTACTCAGAAAT 587


RESULT 14
BH654909
LOCUS       BH654909                807 bp    DNA     linear   GSS 19-FEB-2002
DEFINITION  BOHTZ17TF BO_2_3_KB Brassica oleracea genomic clone BOHTZ17,
            genomic survey sequence.
ACCESSION   BH654909
VERSION     BH654909.1  GI:18713150
KEYWORDS    GSS.
SOURCE      Brassica oleracea
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 807)
  AUTHORS   Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
  TITLE     Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .807
```

```
                     /organism="Brassica oleracea"
                     /mol_type="genomic DNA"
                     /strain="TO1000DH3"
                     /db_xref="taxon:3712"
                     /clone="BOHTZ17"
                     /clone_lib="BO_2_3_KB"
                     /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
                     genomic DNA inserted into pHOS1 using BstXI linkers"
RIGIN

 Query Match           25.8%;  Score 27.6;  DB 28;  Length 807;
 Best Local Similarity  19.0%;  Pred. No. 2.8e+03;
 Matches   19;  Conservative   14;  Mismatches   67;  Indels    0;  Gaps    0;

y          3 AAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNNAAAAANUANAAACAUNNNNNCUU 62
             ||||:      |: :      |:   :::      ||||  :| ||| |:      ::
b        504 AAAATATTTGATATGTATATAGTCTTATTTTTATCACAAAATATAAAAATATTACTAATT 563

y         63 NAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUU 102
              |                |||:                ::
b        564 AAAATTTAAAAATAGTTTAAATTTTATATGAATATAAATT 603


ESULT 15
B050688
OCUS       CB050688                 204 bp    mRNA    linear   EST 17-JAN-2003
EFINITION  NISC_gj19b02.y1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:3272018
           5', mRNA sequence.
CCESSION   CB050688
ERSION     CB050688.1  GI:27788975
EYWORDS    EST.
OURCE      Homo sapiens (human)
 ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
EFERENCE   1  (bases 1 to 204)
 AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
           Tumor Gene Index
 JOURNAL   Unpublished (1997)
OMMENT     Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-r@mail.nih.gov
           cDNA Library Preparation:
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
            DNA Sequencing by: National Institutes of Health Intramural
           Sequencing Center (NISC)
            Clone distribution: NCI-CGAP clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           info@image.llnl.gov
           Plate: LLAM8010  row: D  column: 3
           Seq primer: M13RP1 reverse primer (ABI).
EATURES              Location/Qualifiers
    source           1. .204
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3272018"
                     /sex="male"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Pr28"
                     /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Plasmid DNA from the
                     normalized library NCI_CGAP_Pr22 was prepared, and ss
                     circles were made in vitro. Following HAP purification,
                     this DNA was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from a pool
                     of 5,000 clones made from the same library (cloneIDs
                     985608-986759, 1101192-1101959, and 1217928-1220615).
                     Subtraction by Bento Soares and M. Fatima Bonaldo.  "
RIGIN

 Query Match           25.6%;  Score 27.4;  DB 14;  Length 204;
 Best Local Similarity  24.3%;  Pred. No. 5.3e+03;
 Matches   25;  Conservative    9;  Mismatches   69;  Indels    0;  Gaps    0;

Qy         2 GAAAAUNNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNNAAAAANUANAAACAUNNNNNCU 61
             |||||      : | |         :    :        ||||  :| ||||.:      |:
Db        99 GAAAACTCTTGTTAGACATGATGTTTCCTATACAGAAGAAAAGATACAAACTTGATTGCT 158

Qy        62 UNAGNNNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNUUAG 104
             : ||              ||| |:                : ||
Db       159 TCAGTAAAAACACTTTTAGAGATTGTGTTTCCTTTTTTGTGAG 201


Search completed: March  8, 2004, 06:00:27
Job time : 2387 secs
```